THE DESIGN AND CONDUCT OF THERAPEUTIC TRIALS

A THESIS SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY OF THE UNIVERSITY OF EDINBURGH

- by -

C. C. M. JAMES, M.A.(CANTAB.), F.R.C.S. EDINBURGH

FROM THE DEPARTMENT OF PUBLIC
HEALTH AND SOCIAL MEDICINE,
UNIVERSITY OF EDINBURGH.

PART I. INTRODUCTION.

(i) PRELIMINARY CONSIDERATIONS.

In his inaugural lecture as Regius Professor of Physic at Cambridge University in 1946, Sir Lionel Whitby said: "One has often heard it said that Medicine is an art and not a science. In my own experience, this platitude has usually been a cloak to cover ignorance of some modern scientific procedure. The simple facts are that medicine is both a science and an art. It is true that medicine will never be an exact science, because the normal variations in individuals have such a wide range that automatic and mechanical treatment is prohibited, whilst every patient requires a different method of approach according to his psychology. This, the frequently ridiculed bedside manner, which secures the confidence of the patient, is of fundamental importance in practice; it is an art which is inborn and not acquired. Nevertheless, even the most confirmed artist cannot now neglect to keep up to date with the facts and findings of science upon which his diagnosis and treatment must be based".

With increasing knowledge of aetiology and pathology, empirical methods of treatment have been replaced by those based on laboratory experiment, but before such remedies are used on man, much

preliminary work is needed to ensure that they are not unduly toxic and that they do alleviate or cure. Even though the more poisonous substances are eliminated, it cannot be assumed that because drugs are effective and non-toxic in animals, they will also be in man.

Some drugs have different action on different animals. Morphia, for example, is a depressant in the dog but frequently a deliriant in the cat. Digitalis and others of its group have a variable degree of toxicity according to the species receiving them: Rodents, Toads and the Grass-snake are particularly resistant. The fatal dose for rats is one thousand times greater (per Kilo) than that for cats and thirty times as much as rabbits. Similarly pigeons tolerate intravenously 25 times as much (per unit of body weight) as do cats. In perfusion experiments different reactions occur in the hearts of frogs and toads. Within a species also there is marked variation in response to drugs. De Lind van Wijngaard (1926) who examined the variability of response of 573 cats to a given preparation of Digitalis found that the individual lethal doses were distributed over a wide range. As a result he defined the Mean Lethal Dose since used as a basis for drug standardisation where chemical estimation is not possible.

Since individual variation is so marked, the use of experimental animals necessitates considerable forethought, both as re-

gards the species, for the reasons stated above, and the conditions in which they are used: the Hogben Test for pregnancy using Xenopus was not considered by some investigators (Shapiro and Zwarenstein 1933) to be reliable because they did not understand the care these toads must be given in captivity, but once these difficulties had been overcome (Landgrebe 1939) the test proved simple and trustworthy. One way of decreasing the bias due to individual variations in non-lethal tests, e.g. insulin standardisation in rabbits, is for the control and experimental groups of animals of one days' work to be used next day respectively as experimental and control groups. Before the effect of a drug on an organism in an animal can be tested it is necessary to know something of the habits of the organism and to have available an animal which is susceptible. Until the discovery that the ferret is susceptible to influenza, all research on that infection was held up for lack of a suitable experimental animal. Having found such an animal, the course of the disease must be carefully studied to see that it corresponds with its human counterpart; or what lessons we can learn from differences between the two. Researches on avian malaria have identified part of the life-cycle of the plasmodium not as yet found in man though we have good reason to postulate its existence on other grounds.

The need to validate the relevence of conclusions derived

from laboratory experiments by the proper study of mankind arises also from the fact that the results of *in vitro* experiments are not always confirmed by parallel experiments *in vivo*. In penicillin research, four fractions were originally discovered which had different degrees of anti-bacterial efficacy *in vitro*, the fraction designated K being much more effective than the others, but the K fraction is in the human body so rapidly destroyed that its anti-bacterial action is negligible. In the conduct of experiments of this kind it is essential to eliminate all factors liable to bias results and to make due allowance for experimental inaccuracies which are bound to occur despite the greatest care. These factors and the range of variation from the individual mean can be minimised by maintaining a large number of cases or of experimental animals under standardised conditions.

Whether the subject is man or another species, *Therapeutic Trials* can be defined as "Objective scientific experiments in which are controlled all known variables likely to bias the subjects and objects involved. The purpose of such an experiment being to evaluate a single method, or to compare different methods of treatment". Investigations on man himself are not so simple as those on animals. The circumstances which call for control are more diverse and the co-operation of the experimental subjects is more difficult to obtain especially if the social machinery of medical

care lacks co-ordination. The Army, however, is particularly well suited as a population for experimental purposes. It is about the size of a small nation, but unlike most nations is completely documented with regard to its age and sex distribution and to the social background of its citizens. It possesses an all-in system of social medicine entrusted to a corps of medically trained men responsible to a central authority for the prevention as well as for the cure of disease: furthermore, it maintains a comprehensive system of medical documentation. What is well-nigh impossible in civilian life thus becomes a commonplace, i.e. controlled experimentation on a large scale by pooling the results of many research workers.

It is the object of this thesis to examine Army experience of therapeutic trials. The major factual contribution to this thesis is the section dealing with Syphilis therapy in the Army. For the conduct of this the candidate was wholly responsible, with such clerical assistance as given by junior members of the Army Directorate of Medical (Statistical) Research. The same is true of the sections dealing respectively with Malaria and with Internal Derangement of the Knee. The other sections record collaborative investigations in which the candidate participated as a member of a team of investigators.

(ii) CHOICE OF GROUPS FOR EXPERIMENT AND CONTROL.

In determining the degree of human efficacy of a drug already known to be effective *in vitro* and in animals, it is necessary to study at least two groups of individuals: one treated with the experimental drug, the other as the Control group. The latter are usually treated with another drug whose effects are known. Alternatively, they may receive no treatment other than some neutral substance superficially resembling the drug being tested, as a safeguard against intrusion of the patient's suggestibility. Care in the choice of the two groups is of considerable importance for two principal reasons: (a) they must be of a similar composition with respect to all circumstances known to be relevant or putatively relevant to the prognosis; (b) they should be selected at random to minimise bias arising from unsuspected agencies contributory to the effectiveness of the treatment. Prejudicial selection is most easily avoided when the trial is carefully planned beforehand, being thus an *Ad hoc* Enquiry. Unfortunately it is not always possible to plan a trial beforehand, and much valuable information may be obtained by a *Post hoc* Enquiry.

Ad hoc Enquiries: These may be planned for what may be termed (1) *Confined Populations* in hospitals, barracks or similar communities; (2) *Dispersed Populations*.

(1) Confined Populations. These are communities which are to a considerable extent controllable. They are patients in bed, soldiers in barracks, or children at school. Selection of individuals for experimental and control groups in such communities is most satisfactorily done by allocating individuals in turn to each of the different groups. The individuals may be already in bed, in which case they can be allocated alternately to the different treatments. In out-patients and casualty departments, or in Medical Inspection Rooms in barracks, the patients may be allocated alternately as they arrive. This method of choice is the surest way of obtaining random samples.

When the end in view is the efficacy of preventive measures in the epidemiological field, such alternate selection of individuals cannot be used since it interferes with epidemic spread within the group. The whole group must be alike, and alternate groups must be the basis of comparison. In a trial of Influenza vaccine in Army centres receiving batches of new recruits for training at regular intervals, a whole intake was used as a group; if an epidemic were to develop, the incidence of influenza in a clear minority or a clear majority of the group indicated the efficacy or otherwise of the vaccine, but a middle result might indicate either an avirulent influenzal epidemic, a high level of active immunity among the individuals, or an ineffective vaccine. If the group consisted partly of vaccinated

and partly of non-vaccinated individuals, providing the vaccine were effective, the vaccinated persons would probably not be carriers and therefore the spread would be hindered. The investigators would be left in doubt as to the result since there could be no majority or minority of individuals developing influenza.

(2) Dispersed Populations. These are communities which are not easily controllable in that they are not living or working within the same institution. The method of selection just described, or a version of it, may also be possible here. Army individuals may be selected by their personal numbers which are allocated on enlistment. The terminal digit is usually taken in such cases, e.g. all men with 5 as the last figure, thus avoiding any possible bias due to simultaneous enlistment of men of the same age groups, and hence one block of numbers tends to be allocated to one age group. Among civilians, the method of alternate selection may be applied in alphabetical order: by this means the Jones are equally divided, and the Scotsmen with prefix "Mac", but this point would need to be watched. Similarly, the inhabitants of a street can be taken by alternate houses.

Post hoc Enquiries. In the past the commonest method of analysing results of therapy has been re-examination of individual patients or extraction of information from them by mail after an interval of time. This form of approach may be the only possible

method, but it introduces many sources of bias which have usually been ignored. During the course of years the circumstances attendant on the use of a treatment inevitably alter, the skill and technique of the therapist may change, nursing may be better or worse and carried out in better or worse conditions, materials used may vary in quality and availability, and other conditions may vary pari passu. Death from other causes may remove treated individuals from the census. Others will not reply to communications from the enquirer, some of them dissatisfied with the treatment because it was unsuccessful, or others, themselves cured, unable to see advantage to themselves in responding.

All these possibilities bias the results of a follow-up enquiry and do not necessarily balance out. Often they are unavoidable in a follow-up conducted by an individual clinician; and too often are aggravated by sources of erroneous interpretation due to faulty records or defective analysis. For instance, insufficient attention may be given to the age-distribution of the individuals constituting the experimental group, regardless of the fact that age affects both the likelihood of incurring the condition under investigation and the response to its treatment. Post hoc enquiries which are to accomplish useful results have to take cognizance of all these weaknesses. To this end a method similar to that carried out of recent years in the Army will be possible when civilian hospital

records are organised to make complete statistical information available. The most satisfactory form of follow-up is that in which all cases are personally seen and none are omitted except those which can be satisfactorily accounted for by death from a known cause or by medical reports from other sources. This method has obvious limitations since the patients often have removed to other districts and do not wish to go to the inconvenience of travelling long distances. The Malaria Paludrine trial cited below was followed by post with 100% response, although several requests were needed in some cases to elicit a reply.

In the Army, a system of documentation has been built up which permits easy extraction of information. All medical documents relating to in-patient treatment are filed in a Central Medical File after various details have been abstracted, coded, and transferred to punched cards. The latter are filed separately by diagnosis according to the theatre in which the patient was serving at the time of his admission to hospital. By use of mechanical sorting, the coded details can be listed or tabulated according to the information required. For example, the machines, as well as giving the number of cases of different diseases occurring in a given theatre of war during a stated period, can also provide data with respect to Army number, age, sex, arm of service, date of admission to and discharge from hospital, and disposal on discharge

therefrom of every patient treated for a specified disease or injury. By linking the Army numbers, access can also be obtained to the subsequent Army history of each case through a central card index maintained for every soldier in the British Army. These lists of Army numbers can also be related to the Central Medical File so that the original documents themselves can be consulted. Post hoc enquiries in the Army have thus been carried out with considerable success. Various means have been used to provide control groups, but usually the difficulty has been overcome by contrasting the experimental therapy with one of known efficacy.

(iii) HOMOGENEITY OF EXPERIMENTAL AND CONTROL GROUPS.

Before beginning the analysis of an investigation of the sort under discussion, it is essential to ensure that the groups chosen are homogeneous. Factors biasing results have already been frequently referred to and they must now be dealt with in detail. They may be classified as Personal, Constitutional, Environmental, and Documentary.

(a) Personal Factors: In this context these include Age, Sex, and Marital status. Certain diseases are known to have a differential incidence with regard to each of these, and some means must therefore be available to enable any resulting bias to be eliminated or counter-balanced. The age composition of the Army changed considerably between 1939 and 1945, there being a steady fall in the proportion of very young men and a concomitant decrease in the oldest (over 45) age group. Direct comparison of any two near years was therefore misleading, since the major sources of wastage varied according to which age group predominated, and the emphasis shifted from Tuberculosis to Peptic Ulcers, and from Peptic Ulcers to Bronchitis as the predominance of older men increased. The A.T.S. was at all times a younger population than the male Army personnel, therefore comparison between crude morbidity rates for men and women may give rise to even more erroneous conclusions than comparisons of morbidity

rates for the same class of personnel at different stages of the war. To eliminate the distorting effect of age, all Army figures were age-standardised by a method similar to that used by Registrars-General in their tables of crude and standardised death rates. For instance, among the Army in Britain in 1943 the female-male ratios of the incidence of Peptic Ulcer was 1:10, but after age-standardisation the ratio was strikingly reduced to 1:5. Marital status is also important, particularly with respect to women. A married woman most frequently does not undertake industrial employment and to that extent leads a sheltered life, consequently a group made up of a mixture of married and single women may be biased by a greater or lesser number of industrial workers. Occupation as a factor in therapeutic trials is considered later as an environmental factor.

(b) Constitutional Factors: It is also desirable to pay due regard to Ethnic Origin, Family History, Diathesis and previous Medical History. The ethnic origin of the individuals in the groups is not a major difficulty in the British Isles, but in other countries, especially in the United States of America, different samples of the population may vary considerably with respect to their ethnic composition. This factor is important, not only because of natural immunity to a disease, but also because of immunity acquired by continual contact with endemics and epidemics. An Army enquiry into the health of troops stationed in the Middle East during 1944

did in fact show that the incidence of morbidity and mortality varied widely between troops respectively of European, Asiatic and African stock. For example, the following figures compare morbidity and fatality rates with respect to Tuberculosis (all sites), and Pneumonia among (a) British troops enlisted in Britain, (b) Indian troops, and (c) African native troops enlisted mainly in East Africa:

Morbidity Incidence Ratios:

	Tuberculosis	Pneumonia
Indians:British ...	4.5 : 1	0.9 : 1
Africans : British ...	8.4 : 1	6.1 : 1

Fatality Rates:

	Deaths as % of Cases of Tuberculosis.	Deaths as % of cases of Pneumonia.
British ...	5.8%	1.7%
Indians ...	40.7%	1.5%
Africans ...	22.6%	1.1%

Fatality Ratios (calculated from above):

	Tuberculosis	Pneumonia
Indians : British ...	7.0 : 1	0.9 : 1
Africans : British ...	3.9 : 1	0.6 : 1

In many experiments, the previous medical history of the individuals may not be important but its neglect may sometimes give

rise to misleading results. In investigating the epidemiology of Infective Hepatitis cases of jaundice due to toxic factors such as Virus infection following intravenous arsenic therapy or plasma transfusion will be disclosed only if histories are carefully taken with these points in view. Similarly, family history and Diathesis are issues which need to be considered, e.g. in a survey of Nephritis, a gouty diathesis might be important. In the study of a hereditary condition particular attention must be paid to locality. Recessive characteristics such as Retinitis pigmentosa appear more frequently as a result of inbreeding, and hence may display a differential incidence in more or less isolated communities. If one group in the experiment were to contain an undue preponderance of individuals from, say, the North-East of Scotland, the results of that group might be unfavourably biased.

(c) Environmental Factors: These include Habitat, Income and Occupation. If we use Habitat to describe all the conditions which are responsible for differences between the mode of life of one man and of another, it includes the influence of income and occupation. In this context, we refer more especially to climate, urban and rural dwelling, density of population, local fauna and ease of spread of epidemics. The types of disease prevalent in a locality are to a large extent governed by the climate in so far as certain infective agents are capable of existing only in hot humid

places, while others are rampant only in cold places. Weather conditions affect not only the ability of an organism to survive but also the capacity of the individual to resist infection. Hence, as is well known, the incidence of certain diseases varies with the seasons, for example the Exanthemata all occur at about the same time every year. With decreased rainfall, decreased dilution of contaminants in water increases personal liability to Enteric Fever. In India increased rainfall brings the vibrio of Cholera into the wells of the lowlands and epidemics immediately occur. Temperature also determines the local fauna, and hence the existence of particular vectors of disease. Typhus occurs most commonly in the winter months when people wash less and congregate for warmth, enabling lice to multiply and pass from one person to the next.

Many of these locality effects may be minimised if the individuals have a sufficiently large income: and it is a commonplace that some diseases, particularly Tuberculosis occur most frequently among the lower income groups owing to bad housing conditions, overcrowding, and inability to feed and clothe properly. Occupation also has its attendant hazards by no means circumscribed by such threadbare themes as factory accidents, silicosis, or T.N.T. toxicosis. For instance the worker may be involved in travelling long distances from his home, or compelled by the nature of his work to irregular habits of feeding.

(d) Documentary Factors: In the Army, the difficulties stated above can usually be overcome: but another remains. In Britain service patients were treated throughout the war either in Military Hospitals or in E.M.S. Hospitals under the control of the Ministry of Health. Treatment and documentation in the two types of hospital was similar but not always identical. Patients were sent to either type of hospital in accordance with policy based on the number of beds, specialists etc. available, and the policy pursued by the Army altered with changing circumstances. In therapeutic trials quoted later, viz., those relating to Impetigo and Internal Derangement of the Knee, attention is drawn to a striking difference with respect to duration of stay in hospital of patients respectively treated in Military and in E.M.S. hospitals. Thus sample populations are strictly comparable only if drawn from comparable institutions.

It is evident that uniform documentation must be established to carry out a therapeutic trial involving different therapeutic centres. An enquiry into the Efficacy of Penicillin in War Wounds recently carried out at the War Office for the Medical Research Council was in fact jeopardised by oversight in this connexion. The experimental group consisted of casualties in North-West Europe, while the control group consisted of casualties in Italy. Owing to differences of policy with respect to documentation of different medical units the experimental group was weighted by a far greater proportion

of very seriously wounded cases, and in consequence included many more deaths. Uniform documentation involves not only use of the same types of document and their complete availability, but also uniformity in completing the document so that the answers allow of no ambiguity either in matter of fact or degree.

In attaining homogeneity of samples in the many ways referred to above, it must also be stressed that no catalogue of avoidable errors is exhaustive. While all known factors likely to bias the selection may have been eliminated or allowed for in planning the experiment, there may always be others which could not be foreseen. The only safeguard against such is the use of artifices of randomisation, as for instance, alternate selection cited in a previous paragraph.

MALARIA INQUIRY

Hospital

1. NAME.....

2. ARMY NUMBER.....

3. RANK .....

4. REGIMENT or CORPS.....5. UNIT .....

6. LENGTH OF EMBODIED SERVICE.....years.....months.....

7. OVERSEAS SERVICE.....

THEATRE	Dates	
	From	To

8. PRESENT ATTACK WARD SERIAL NO.

Diagnosis B.T. M.T. Ovale

Treatment M. 4888 QUININE-PAMAQUIN MEPACRINE OTHER _____

Date of Admission.....Date of Discharge.....

Date of beginning of treatment..... Date of end of treatment.....

Smear before commencement of treatment + - Smear on discharge + -

Date of onset of symptoms.....

Date of beginning of fever.....

Duration of fever after beginning of treatment.....days

HOME ADDRESS A & S. GROUP U.K. DATE OF DISEMBARKATION

9. HISTORY OF PREVIOUS ATTACKS

Total number of prior attacks.....

Details of Prior Attacks

ATTACK	PARASITES	DATE OF ONSET	TREATMENT
1st Attack			
Attacks during past 6 months			

ABBREVIATE TREATMENT Quinine Q
Pamaquin P
Mepacrine M
M.4430)
M.4888) Write in full
&c)

10. LAST DATE OF TAKING ANTI-MALARIAL DRUGS PRIOR TO ADMISSION .....NATURE AND AMOUNT OF DRUG IF KNOWN.....

11. ILL-EFFECTS

SYMPTOM	SEVERITY	DURATION

SUBSEQUENT HISTORY For W.O. use only

12. DATE OF ADMISSION FOR RELAPSE.....

13. DIAGNOSIS: Smear positive B.T. M.T.
Smear negative Q Ovale
Clinical

Fig.1

CONORRHOEA - U.K.

Case No. [ ]

No..... Age..... Diagnosis

Date of commencement of treatment D M Y [ ][ ][ ]

G.C. | URETHRITIS

TREATMENT

Form of Treatment	Period in days from commencement of treatment			Total Dosage
Penicillin				
Sulphathiazole				
Sulphadiazine				
Sulphapyridine				
Sulphanilamide				
T.A.B.				
G.C.Vaccine				
Others				
$KMnO_4$				
HgOCN				

INTERVAL.....days from onset to commencement of treatment.

INTERVAL.....days duration of treatment.

COMPLICATIONS

Conjunctivitis Iritis Arthritis Epididymo-orchitis.
Prostatitis. Vesiculitis Others None

Interval from onset.....days

JAUNDICE Present Absent
Interval start of treatment to onset of Jaundice.....days

RELAPSES

Interval.....days from completion of treatment to relapse

TREATMENT RELAPSED Responded / Failed / Relapsed

PREVIOUS HISTORY

No.of previous attacks (or relapses).....days Not specified
Interval from completion of last attack.....days Not specified

Treatment:- Sulphathiazole HgOCN T.A.B.
Sulphanilamide P.P. G.C.
Sulphadiazine Others
Sulphapyridine

Penicillin Not specified

Fig.2

(iv) DESIGN OF DOCUMENTS.

In both ad hoc and post hoc Enquiries some means must be found to attain uniformity with respect to provision and abstraction of information. The simplest method is to design a document specifically for the purpose. Such a proforma should record only details relevant to an investigation, set out in such a way that the eye is able to perceive at a glance what it is looking for. It can be designed either for ultimate sorting by hand or for abstraction on to the punched card for mechanical tabulation. If designed for hand sorting, the data on the proforma may best be listed either for the relevant word to be underlined or with a box or space alongside for the insertion of a mark such as a cross or a figure; each item being mutually exclusive to avoid ambiguity and to facilitate numerical checks. Every section must therefore include an item for information not specified, and an item for "other", to cover all possibilities: by this means no case in unaccounted for in any section. If mechanical tabulation is envisaged the lay-out is similar but each item must bear a number, and the recorder will mark the relevant number in a box printed on the right-hand edge of the page. This method allows a punch-clerk to transfer the code number directly to a punch-card and avoids the intermediate step of employing a clerk, unskilled in scientific

MEDICAL CODE SLIP (Hospital Cases)

1	THEATRE IN WHICH TREATED	
2	PERSONAL OR ARMY NUMBER	
3	DISCRIMINATION OF PERSONNEL	
4	ARM OF SERVICE	
5	AGE (in years)	
6	TYPE OF HOSPITAL OF 1st ADMISSION	
7	TYPES OF MEDICAL UNIT	
8	MONTH AND YEAR OF 1st ADMISSION (from A.F. W.3118, if available)	
9	DURATION OF STAY IN FORWARD MEDICAL UNIT OR RECEPTION STATION (from A.F. W.3118)	
10	DURATION OF STAY IN HOSPITAL	
11	DURATION OF STAY IN CONVALESCENT DEPOT	
12	DURATION OF STAY - TOTAL (9+10+11)	
13	RESULT ON DISCHARGE	
14	DISEASE OR INJURY	
15	E.A. OR NOT	
16	WEAPON OR AGENCY OF INJURY	
17	NEW DISEASE SUPERVENING OR 2nd DISEASE/INJURY	
18	TREATMENT	
19	MARITAL STATUS	
20	TYPE OF ENGAGEMENT	
21	PERIOD OF SERVICE IN THEATRE (in months)	
22	TYPE OF ARMY DOMICILE	
23	SITE OF INJURY	
24	TISSUES INVOLVED	
	CODED BY	CHECKED BY

Fig.3

terminology, to interpret the information in the process of coding it, a step which increases the chance of mistakes.

Mechanical tabulation is not worthwhile if the number of cases is small, it finds its greatest value in organizations where records are always maintained with this end in view. Hand sorting may be very laborious and may take a long time, but is more flexible than mechanical tabulation, and may disclose information which would otherwise escape notice. For hand sorting it is important that the paper or card on which the proforma is printed should be semi-stiff, neither thick nor very thin. Specimen proformas which have been used in various enquiries designed for hand sorting are illustrated in Figures 1 and 2. Figure 3 shows an Army form for mechanical tabulation, but necessitates the employment of clerks skilled in coding scientific terms.

(v) CRITERIA OF EFFICACY.

Ideally, estimation of the effect of a therapy should rely upon exactly measurable criteria of which there are very few in the human field. It should eschew personal impressions which may have no objective validity and are difficult or impossible to check. We have thus to ask: What objective criteria are available?

The course of any disease terminates with cure or with death. Which of these will happen depends upon the disease itself, the effectiveness of the treatment and time factor. In the absence of treatment, diseases may be classified as follows:-

1. always fatal.
2. always cured without complications.
3. usually cured but otherwise fatal: sometimes accompanied by complications.
4. becoming chronic.
5. subject to relapses.

Under treatment the outcome may be any of the following:-

1. The course of the disease may be unaffected.
2. The patients may survive in contrariety to previous experience.
3. Cure may be absolute, without complication or relapse.
4. Complications may be prevented either in frequency or in severity.

5. Complications may occur in a previously fatal disease, without death following.

6. The risk of relapses may be reduced.

7. There may be variation in the time factor in relation to each of the foregoing, e.g. death may occur at a later stage of the disease.

8. The treatment may maintain the status quo, e.g. Insulin in Diabetes.

A more detailed examination of the above suffices to indicate possible yardsticks of therapeutic efficacy.

(a) Mortality Rate. We may consider the risk of death with respect to a disease as (1) immediate, (2) delayed. Accordingly we may use as our criteria of efficacy one of two indices, as circumstances indicate.

(1) The Immediate Mortality Rate may be defined as that occurring within a period after treatment during which a level of response is capable of being attained, i.e. the period up to discharge from hospital. A patient is not moved from hospital until he has recovered sufficiently to permit him being moved and only very rarely if he is expected to die within a short period. The implication in utilisation of this period is that the patient has recovered to a measurable degree. It would be exceptional for a patient to be discharged home without attaining this degree of recovery. The actual duration of the period cannot be taken as a measure since it depends so very much more upon the individual patient and his powers of re-

cuperation than upon the efficiency of the therapy. In any comparison using this yardstick, care must be taken to ensure that administrative policy in relation to such classes of patients was the same at each hospital and throughout the period.

(2) Life Expectation: The expectation of more or less prolonged life is in any circumstances dependant on the age of the individual, hence the only satisfactory measure of delayed risk of death is a comparison of the expectation of life of treated individuals at a given age with that of untreated individuals of the same age, in accordance with the actuarial method of the Life Table. The same method is adaptable to assessment of the risk of relapse as illustrated by the section below dealing with Syphilis Therapy.

(b) Relapses. What is a relapse and what is not a relapse can usually be distinguished, but in such conditions as Gonorrhoea, Syphilis, Malaria and Infective Hepatitis the distinction between a relapse and a re-infection is not always easy to make. Signs and symptoms of re-infection and relapse are not easily separable: and their separation may involve an act of personal judgment, the value of which will depend on the experience of the investigator. Criteria of Relapse used in any trial must therefore be stated with precision to permit comparison with subsequent trials. This question is dealt with in the trials of Gonorrhoea and Syphilis therapy reported later.

(c) Duration of Treatment: In diseases which do not usually have a fatal termination, treatment may be able to reduce the length of time during which the sufferer is incapacitated. Duration of treatment is therefore a possible yardstick of its value. Gonorrhoea and Impetigo trials discussed below illustrate its use and its limitations. Its limitations vis-à-vis hospital administration, already mentioned in relation to documentation, are the subject of further comment in the Section on Internal Derangement of the Knee. This criterion can be reliable only when patients are received in rotation in the same clinic and treated by the same doctors, nurses and administration; and should not be used when prolongation of treatment is due to the occurrence of complications, which themselves provide a criterion of efficacy.

(d) Supervening Diseases and Complications: This yardstick can conveniently be used in situations such as those which prescribe recourse to mean duration of hospitalisation. The two may in fact be estimates of the same thing in different ways, although it is obviously inaccurate to judge different therapies by the speed with which they overcome the primary condition when this duration may be grossly prolonged by the occurrence of a secondary condition or complication. In estimating the efficacy of Penicillin prophylaxis in Battle Casualties, the incidence of infective complications proved to be the only satisfactory criterion. Another ex-

ample of its applicability is comparison between frequencies of chest complications associated with use of different anaesthetics for a single type of operation performed on similar subjects.

(e) Vocational Efficiency: Discharge from the Army in Category E (medically unfit for further service) constitutes a clear-cut measure of the severity of certain diseases after attempted cure, and has its industrial parallel in disablement. Any data which take account of personal efficiency, such as output per man-hour, rates of absenteeism, rates of promotion or increased pay can also be used to assess the effectiveness of previous treatment for a more or less disabling disease. Some criteria available in the Army are not available in civil life. Since every man and woman has a medical examination on enlistment and is assessed by medical category, change in the category necessitated by disease or injury provides information relevant to our present purpose: but there are many difficulties in the way of this method. Comparison of series of cases collected over different periods may contain populations different with respect to their medical category composition. Army policy with reference to medical categorisation changed from time to time according to the man-power situation. At the time of greatest shortage low category men were upgraded to A1 or sent to work in places where it had never previously been considered right for them to work. Where policy changes have not occurred, then, change of medical category is a useful yardstick.

In the section on Internal Derangement of the Knee (I.D.K.) the subject of downgrading and discharge from the service is illustrated, and attention is drawn to the misleading results obtained if administrative factors are overlooked. In this context, in view of the frequency of football as a causative trauma a yardstick of efficacy could have been the rate of return to full work of professional footballers after different methods of therapy.

In choosing a satisfactory criterion, the opportunity of using additional secondary criteria should not be missed. Such secondary, or even equivocal criteria not only provide extra evidence, but also act as internal checks one upon the other. For the Therapeutic Trial of Gonorrhoea reported below three yardsticks were available. Where different therapies are used as standards for each other, it is advisable to have one established method as a basis for comparison.

(f) CURE.

Since the body is continually in a state of repairing damage, our criterion of disease is not clear-cut. All damaging processes call further defensive and repair efforts by the tissues and it is the degree of response required which governs the designation of a particular attack as a disease. A disease may therefore be defined as an injurious process by foreign organisms which necessitates more than a purely local defensive and repair effort by the body. Cure of a disease may be defined as the point at which a disease process is terminated, whether the body is or is not severely disabled by it. In this sense of the term Cure has no relation to the degree of resultant disability. After Broncho-pneumonia, the amount of fibrosis occurring in the repair process will determine the degree of disability, which persists.

Whether such a definition of cure is or is not en rapport with universal usage, it provides a pigeon-hole for the criterion we commonly employ, when we seek to evaluate a method of treatment applicable to lingering diseases (e.g. Eczema) which are not themselves disabling, nor fatal nor prone to relapse after disappearance of the signs and/or symptoms. The arrest of the disease process as a criterion of remedial efficacy itself presupposes that we have the means of recognising it as such. Only if the differential diagnosis of a disease is clear-cut can we use cure as defined above to assess

the measures we employ in the patient's interest.

In the practice of medicine certain conditions with clear-cut symptoms and signs which run a definite course occur with sufficient severity and sufficiently frequently among the population to justify their denomination as specific entities. Where symptoms, signs and the course of a disease are less specific the synchronism or sequence may itself be sufficiently specific to rank as a characteristic syndrome: but it is by no means true that all conditions which have earned a name are singular entities either from a symptomatological or an aetiological viewpoint. When doubt exists about a correct diagnosis the physician may turn to a therapy designed for a specific disease as a diagnostic measure. In other words conversely to testing the efficacy of a therapy in a known disease. The delimitation of the disease and the definition of cure became entangled for another reason. A particular microorganism may be classifiable in strains distinguished by no recognisable criteria other than their resistance to a particular drug, and the symptomatology of the disease associated with one or the other may in fact admit no differential diagnosis, as is true of sulpha-fast and sulpha-sensitive cases of Gonorrhoea. From a clinical viewpoint we are here dealing with one disease which responds erratically to a specific treatment. From an aetiological viewpoint we might with propriety speak of two similar diseases for one of which the same treatment is completely

efficacious.

In comparing results of therapy by recourse to the criterion of cure as here defined, it is necessary to pay regard to the possibility that the disease process could arrest without interference. It may therefore be necessary to define (1) the point at which cure is assumed to occur, (2) the indications for determining this point, and (3) the reason for accepting the indications and the timing as being the time of cure. However, more often than not, the yardstick of cure is used only when it is necessary to state the condition was cured within a given period: the exact moment at which it occurred does not matter.

(vi) SUMMARY OF PART I.

Throughout this thesis, there has been constant reference to elimination of bias and the avoidance of 'selection'. The intention of a therapeutic trial is to obtain reliable information as to the efficacy of a drug. In designing the trial every consideration must be given to avoiding anything which will prevent an exact comparison between different methods of therapy. It is essential to take stock of all factors which can be considered to be of possible importance and to control them. Absolute homogeneity is impossible of realization, but suitable devices for randomisation of test subjects should suffice to minimise bias due to unknown or unsuspected agencies in a large-scale investigation. What constitutes a large-scale investigation must be considered with due regard to the magnitude of the difference an experiment is designed to detect. If we are comparing two methods of treatment either of which can guarantee at least 90% efficacy, the standard error of a sample of 900 is as great as 1%. We could not judge samples of 1000 each as certainly different unless they differed by more than 3%. A therapeutic trial must be regarded as a test of a procedure only in the conditions and circumstances of the trial. In the trial of Paludrine as curative treatment for B.T. Malaria, given below, the high relapse rates recorded, in the neighbourhood of 50%, might tempt one to conclude that Paludrine is totally ineffective as an anti-

malarial drug. Other workers, using low dosage, but for a longer period of time, have in fact shown Paludrine to be an effective cure of B.T. Malaria.

Post hoc statistical techniques to test for homogeneity of experimental data are sufficiently dealt with in theoretical treatises accessible to the research worker, but it is permissible to doubt whether such devices can suffice to safeguard against errors which first-hand experience of the subject-matter might anticipate. In this context we may recall Bacon's aphorism that radical errors in the first concoction cannot be cured by subsequent remedies however excellent. It is the aim of this Thesis to illustrate how safeguards against erroneous interpretation emerge from the practical conduct of Therapeutic Trials and to show by an examination of the credentials of particular therapies how any investigation of this sort is beset by pitfalls peculiar to itself.

PART II: ARMY EXPERIENCE OF SYPHILIS TREATMENT.

1. INTRODUCTION.

Experience of Syphilis therapy in the Army during the late war provides unique materials for assessing therapies in current use, and it is the aim of Part II of this thesis to assess their value. During the first phase of the war Army medicine relied exclusively on long-term arsenotherapy, later on short-term Mapharside with Bismuth. In the last phase Penicillin treatment replaced or supplemented the use of Arsenicals for a trial period pending the outcome of experience. In all, we are in a position to examine the results of eight procedures which we specify more fully below (§ 5):

1. 2.4 mega units of Penicillin.
2. 4.0 mega units of Penicillin.
3. 2.4 mega units of Penicillin + 0.6 gm. Mapharside.
4. 2.4 mega units of Penicillin + 0.4 gm. Mapharside + 1.0 gm. Bismuth.
5. Long-term Arsenic Therapy high aggregate dosage.
6. Do. Do. low aggregate dosage.
7. Long-term Arsenic Therapy + 2.4 Penicillin.
8. Short-term Arsenic Therapy.

2. SOURCES OF MATERIAL.

Long-term Arsenic Cases were traced through the War Office Central Card Index which supplied a nominal roll of all Syphilis cases admitted to hospital for treatment during the twelve months 1943-44. The same nominal roll also served for a survey of post-arsphenamine jaundice (Truelove and Hogben, 1947). Nominal rolls obtained from the Adviser in Venereology made it possible to trace cases undergoing initial treatment (by methods later specified) at Army Special Treatment Centres in Britain during the period September 1944 to June 1946. During this period each Special Treatment Centre (S.T.C.) was treating all cases of early Syphilis with a standard schedule of therapy notwithstanding the duration or severity of the disease. Hence such nominal rolls should provide an unbiased sample of the total Army early Syphilitic population at home. By recourse to the serial numbers supplied by them, it was possible to extract hospital record cards (A.Fs. I.1220) from the Central File of Medical Documents at the War Office and from the Central Syphilis Register (C.S.R.). The former furnished detailed information with reference to initial syphilitic infection and treatment. The latter supplied surveillance notes relating to each case.

Many cases on the nominal rolls from S.T.Cs. were unserviceable for the following reasons:-

(1) patients from other Services, initially treated in military units were subsequently transferred from Army supervision;

(2) some patients had had previous treatment for the disease;

(3) some cases were congenital, tertiary, neuro-syphilitics or latent syphilitics;

(4) the relevant documents of a small group were incomplete;

(5) a large group of cases (here termed Defaulters) were demobilised too soon after treatment to permit surveillance. This circumstance is responsible for the disappointing yield of cases treated by the more recent Penicillin Therapy Schedules, but there is no reason to suppose that their exclusion gives rise to bias.

All cases under consideration are Early. They had neither suffered previously from the disease nor undergone previous treatment, but otherwise include any diagnosed as Sero-negative Primary Syphilis, Sero-positive Primary Syphilis and Early Secondary Syphilis. The diagnostic criterion of Primary cases was presence of a primary chancre confirmed by detection of Treponema pallidum by dark ground examination, with or without positive serological test (Wassermann or Kahn). That of Secondary Syphilis was a positive Wassermann or Kahn Test together with such clinical manifestations as a rash, mucous patches or condylomata (confirmed as such by dark ground examination).

3. CRITERIA OF EFFICACIOUS TREATMENT.

Inter alia we may judge the therapeutic efficacy of a drug in terms of the risk of relapse, or the frequency of beneficial response whether temporary or otherwise. Accordingly, we here distinguish two categories of cases as relapses and failures, the former among whom the benefit of treatment is only temporary, the latter because they fail to derive any manifest benefit at all. Any index which combines figures referable to both criteria is necessarily arbitrary, because we have no reason to suppose that a treatment which is accredited by a lower relapse rate will be accredited by a lower failure rate: and if the two methods of assessment point to different conclusions, no formula applicable to all diseases prescribes which should have priority. Since we cannot compute relapse rates without excluding cases which at no time responded to treatment, we have thus to apply to our data two independent criteria of efficacy: (a) what proportion of all cases treated fail to respond at any time; (b) among cases which do respond at some time, what proportion subsequently relapse.

Considerations which did in fact guide our classification of cases as failures, relapses or cures, call for more explicit clarification. Both for diagnosis and for surveillance, Treatment Centres concerned in the present investigation employed the Quali-

tative Wassermann reaction (W.R.), the Qualitative Kahn Test (K) and Quantitative ditto (K.Q.). Owing to exigencies of the Service, the W.R. and Kahn were not always used together in surveillance; but Quantitative Kahn tests (K.Q.) performed in the majority of cases give evidence of serological trends. We here consider a serum to be positive when the laboratory reported any one of the following: W.R.+, K+ or K.Q. of 4 or more units. We classify as doubtful a result such as K±, W.R.± or K.Q. readings of less than 4 units. Blood samples from each case were commonly examined both at beginning and end of treatment. The routine of surveillance entailed clinical and serological examination at 2, 4, 6, 9, 12, 15, 18 and 24 months after completion of treatment. Lumbar puncture was performed at 6, 18 and 24 months.

We here regard a case as cured, if it conforms to each of two criteria: (a) all lesions have healed at the conclusion of treatment; (b) the serological test is negative at the end of surveillance. We also regard as a cure, any case ending surveillance prematurely with one sero-negative test (e.g. W.R.) and one doubtful (e.g. Kahn) if it satisfies each of the following criteria: (a) there has been consistent decline from an originally high titre; (b) there are no remaining clinical signs of syphilis. Clinical records (A.Fs. I.1220) provide the information on which we rely for specification of relapse in accordance with either of the two criteria

on which we base the initial diagnosis, viz: (a) presence of fresh clinical manifestations in which Treponema pallidum could be demonstrated; (b) positive serological test, if previous tests had been negative or of lower titre. Reappearance of Muco-cutaneous lesions in which T.p. is demonstrable provides clear-cut evidence of relapse; but the serological criterion is not, of itself, unequivocal.

Many cases lost in the follow-up (Defaulters) were not sero-negative at the time. They were either sero-positive or serologically doubtful; and their previous serology may have been: (i) positive with a steady or rising titre; (ii) positive with a declining titre; (iii) consistently or at some time negative. Cases we classify as failures, in contradistinction to relapses, are cases which did not become sero-negative within four months of the end of treatment but continued with a high titre or an increasing titre until they were lost to surveillance. In this context failure therefore means inefficacy of a treatment to reverse or to reduce the serological titre by the fourth month. There were few among the long-term arsenic series, and only among such as received small aggregate doses of arsenic owing to supervening dermatitis or jaundice. In the Penicillin and Short-term Arsenic series some failures renewed treatment as soon as there was clear evidence of defective response. The rest would almost

certainly have been re-treated, had they not defaulted.

Cases defaulting from surveillance with a titre still positive but steadily declining from an initially high level are obviously doubtful. Cases with a positive serum after one or more clear-cut negative tests during surveillance are by definition Relapses at the month in which they become positive. Others classified as doubtful had a doubtful serum reaction at their last surveillance before they defaulted, because further surveillance might have demonstrated a steadily rising titre or fresh muco-cutaneous lesions indicative of relapse. In a very few instances, a weakly positive serum reaction interrupted an otherwise consistently negative surveillance. The subsequent negative clinical and serological trend, confirmed by absence of abnormality in the C.S.F., suggested that such isolated positive reactions were either of the falsely positive type or were due to laboratory errors. Little can be predicted of a case which defaults from surveillance at the 2nd or 3rd month with a serum still positive or doubtful and without any stigmata of clinical relapse. Since the time under surveillance is quite inadequate to disclose the ultimate trend, we have rejected them from our pool of data. Table I shows the percentage of all cases available for follow-up in each series and those rejected for reasons here stated.

TABLE I.

Percentage of cases ending with an equivocal serology rejected from the series on account of follow-up of less than four months.

	PRIMARY SERO-POSITIVE		SECONDARY	
	Total Cases	% Cases rejected	Total Cases	% Cases rejected
PEN: 2.4	163	0.0	119	0.8
PEN: 4.0	44	6.8	30	6.7
PEN: & MAPH:	113	2.7	59	10.2
PEN: & MAPH: + BISMUTH	92	3.3	56	14.3
LTA: & PEN:	38	0.0	26	0.0
LTA: (A)	84	0.0	67	0.0
LTA: (B)	84	0.0	47	2.1
STA:	38	0.0	42	0.0

In Therapeutic Trials involving a communicable disease a common difficulty arises in connexion with delimitation of relapses and re-infections. Eighty-five cases in all treatment groups of our series were known to develop additional syphilitic

lesions. The venereologist's opinion as stated in the clinical records distributes the later lesions as follows:-

Relapses	33
Reinfections	12
Inconclusive	40

It has hitherto been customary (Moore 1943) to exclude a diagnosis of reinfection unless consonant with all of the three following criteria:

(i) "There must be proof that the patient had syphilis prior to the occurrence of the suspected second infection, and this proof must rest on the demonstration of spirochetes in a lesion or the occurrence of a positive blood Wassermann reaction in the blood serum, and not on clinical judgment alone".

(ii) "After an interval following antisyphilitic treatment and at a site other than that of the primary lesion of the first infection, there must develop a lesion with the characteristics of a chancre in which spirochetes can be demonstrated".

(iii) "That at the time of the supposed second infection the serologic test is negative; and that while under observation and during a period in which treatment is purposely withheld, the serologic test becomes positive or the patient develops outspoken secondary syphilis".

The clinical records of our cases do not indicate that the venereologist's opinion relied on the coincidence of these three criteria. The first of Moore's criteria (i) was indeed fulfilled in all our cases; but analysis of the sample on the basis of the other two yielded the following figures:

	Sore at original site		Sore at new site		TOTAL
	Seropositive	Seronegative	Seropos.	Seroneg.	
Cited as relapse.	28	3	2	0	33
Cited as reinfection.	3	2	6	1	12
Non-committal	26	7	3	4	40
TOTAL:	57	12	11	5	85

Records usually laid emphasis on a history of recent exposure where citing a diagnosis of reinfection. It is reasonable to assume that the 57 cases whose second lesions appeared at the site of the original chancre with coincident positive serum reaction were definite cases of muco-cutaneous relapse. Probably, the 12 cases which were seronegative when a second lesion appeared on the original

site were also muco-cutaneous relapses. Inter alia, Hudelo and Rabut described 21 such cases in a series of 51. Muco-cutaneous relapse in the presence of a negative or doubtful serum is thus well accredited. The 11 cases whose second lesions were stated to have occurred at different sites from the original chancres at a time when the serum was positive are more difficult to assess. It was clear from the notes of some that the new lesions were sufficiently close to be within the lymphatic watershed of the original site, a fact which might well suggest that they were probably muco-cutaneous phenomena. On the other hand, records of one or two with typical chancres proved to contain T.p. at a site well away from the original chancre and associated with enlargement of lymph glands other than those stated to be enlarged in the original infection cite a doubtful serum reaction. It is conceivable that such cases were, in fact, reinfections. Of the five remaining cases, none fully satisfies the criteria advanced by Moore. We therefore regard them as relapses, and cite further details with reference to them in Table II.

TABLE II.

Case	Age	Site of first sore	Sites of Adeno-pathy	Initial Serology	Treatment	Serology after treatment to time of devel-opment of 2nd sore	Time of appear-ance of 2nd sore after treatment of 1st.	History of new exposure to in-fection	Site of 2nd sore	Sites of Adeno-pathy.	Serology at time of devel-opment of 2nd sore	Vener-eologist's Diagnosis
1	42	Left side of glans and sulcus DG Tp +	Inguin-al Bilat-eral	(W.R.not done) Kahn neg-ative	PEN: 2.4 Mega units	Negative at sec-ond month	Third month	3/52 E.M.C. 6 days marital	The Meatus Tp +	Inguin-al Bilat-eral	(W.R.not done) Kahn negative	"Syphilis Early"
2	29	Coronal Sulcus DG Tp +	Inguin-al Bilat-eral	W.R. & Kahn negative	PEN: 2.4 Mega units	Consist-ently negative	Eight-eenth month	2/12 E.M.C. 3 days marital	Coronal Sulcus DG Tp + (not on original site)	Inguin-al left	(W.R.not done) Kahn negative	"Syphilis Early"
3	20	Prepuce and glans DG Tp +	Inguin-al Bilat-eral	W.R. & Kahn negat-ive	PEN: 2.4 Mega units + Maph: 0.6 gm.	Consist-ently negative	Sixth month	6/52 Prostit-ute	Papule in Rt. Inguinal region DG Tp + Penis clear	No Adeno-pathy	(W.R. not done) Kahn neg-ative	"Syphilis Primary"
4	33	Fraenum DG Tp +	Inguin-al Bilat-eral	(W.R.not done) Kahn Positive	PEN: 2.4 Mega units	Negative at 2nd month	Third month	Not stated	Prepuce DG Tp +	Inguin-al Bilat-eral	(W.R.not done) Kahn Negative	"Reinfec-tion"
5	25	Ventral surface of shaft of penis DG Tp +	Inguin-al left	(W.R.not done) Kahn doubt-fully positive	Mapharside & Bismuth (S.T.A.)	Consist-ently Negative	Seven-teenth month	2/12 E.M.C.	Dorsum of Penis DG Tp neg: 1/12 after sore appeared	Not stated	Kahn neg: when sore appeared. Became pos-itive be-fore re-treatment	"Syphilis Early"

(E.M.C. = extra-marital coitus)

The evidence that the second sore differed significantly in respect of site from the initial chancre is poor in cases 1, 2 and 4. In case 3 the new lesion was apparently still within the lymphatic watershed of the initial chancre. Case 5 is the only one in which a negative serum reaction in the presence of the second sore was observed to have become positive before further treatment was instituted. Unfortunately, dark ground examination was not done until the sore had been present for one month. It was then negative on three occasions. In no case was there evidence that 'repeat' blood tests had been used to confirm routine tests.

4. CALCULATION OF RELAPSE RATES.

All time periods date from the end of the course of treatment. In what follows, the assumption is that we have sorted our data to exhibit:

(a) numbers of seronegative individuals (x_{m-1}) still under observation at the beginning of the mth month from date of certification;

(b) ditto (x_m) still remaining under observation at the end of the mth month;

(c) relapses (y_m) during the mth month;

(d) defaulters (d_m) i.e. individuals who have ceased to be under observation during the mth month.

The last item is evidently deducible from the others by the relation:

$$d_m = x_{m-1} - x_m - y_m \quad \ldots\ldots\ldots\ldots \text{(i)}$$

Our problem is to ascertain the chance of relapse during the mth month, denoted below as q_m. If there were no defaulters, this would be the ratio of relapses during the month to seronegative individuals at the beginning, i.e. $y_m \div x_{m-1}$; but d_m defaulters would have accounted for an unascertained number (e_m) of relapses during the same period, so that the correct value is given by:

$$q_m = \frac{y_m + e_m}{x_{m-1}} \quad \ldots\ldots\ldots\ldots\ldots\ldots \text{(ii)}$$

Our problem is therefore to estimate e_m. Let us suppose that all the d_m individuals had ceased to be available for further observation on the first day of the mth month after reporting for the census recording x_{m-1} seronegative individuals at that date. If so, e_m would be the relapses during the month referable to an initial population of d_m seronegative individuals, i.e.

$$q_m = \frac{e_m}{d_m} \quad \text{or} \quad e_m = q_m d_m$$

In fact, we are entitled to assume that the stream of defaulters is steady, the mean time during which they were available for observation being $\frac{1}{2}$ a month. We may therefore put approximately

$$e_m = \tfrac{1}{2} q_m d_m$$

It is important to point out that there is only one arbitrary assumption in the above relation, that is to say, a linear rate of default in the interval. Otherwise the relation is exact for the following reason. Suppose that we divide the month into a large number (s) of intervals in each of which the number of defaulters is $d_m \div s$. At the beginning of each interval the number of defaulters is thus the same; but their contributions to the total (e_m) of unidentified relapses will

be different. The first batch with a month before them can contribute to this total $q_m\ d_m \div s$. The second batch with a residual fraction of a month equivalent to $(s-1) \div s$ can contribute only $\overline{s-1}\ q_m\ d_m \div s^2$ and so on. The best estimate of e_m is the limiting value of the corresponding sum of the contributions of the batches when $\underline{s}$ is indefinitely large, i.e.

$$e_m = \text{Lt}\ \frac{q_m\ d_m}{s^2} \left(s + \overline{s-1} + \overline{s-2} \ \ldots\ 2 + 1 \right)$$

$$\therefore\ e_m = \tfrac{1}{2}\ q_m\ d_m.$$

Hence by (i) above

$$e_m = \tfrac{1}{2}\ (x_{m-1} - x_m - y_m)\ .\ q_m \qquad \ldots\ldots\ldots\ldots\ldots\ (iii)$$

Likewise by (ii) above

$$q_m = \frac{y_m + \tfrac{1}{2} q_m\ (x_{m-1} - x_m - y_m)}{x_m - 1} \qquad \ldots\ldots\ldots\ldots\ (iv)$$

$$\therefore\ 2\ q_m\ .\ x_{m-1} = 2y_m + q_m\ x_{m-1} - q_m\ x_m - q_m\ y_m$$

$$\therefore\ q_m\ (x_{m-1} + x_m + y_m) = 2y_m$$

$$\therefore\ q_m = \frac{2y_m}{x_{m-1} + z_m + y_m} \qquad \ldots\ldots\ldots\ldots\ (v)$$

The appropriate computing scheme for the estimation of q_m is thus:

Date by month (1st day of)	Residual seronegative population	Serological relapses.	Defaulters from equation (i)	q_m from (v)
0	x_o			
.....		y_1	d_1	q_1
1	x_1			
.....		y_2	d_2	q_2
2	x_2			
.....		y_3	d_3	q_3
3	x_3			

Corresponding to q_m the relapse rate, the survival rate s_m showing the chance of remaining negative throughout the month m is given by:

$$s_m = 1 - q_m$$

The chance of remaining seronegative to the end of the mth month from the beginning of the follow-up is the combined product

$$\prod_{1}^{m} s_r$$

With a computing machine the effort involved in extracting the continued product is small; but it is not a necessary operation and the two methods of arriving at a result should be the same. The second is to proceed thus:

Residual seronegatives	Observed Relapses	Estimated Relapses	Cumulative estimated relapses
x_0			
	y_1	$y_1 + e_1$	$y_1 + e_1$
x_1			
	y_2	$y_2 + e_2$	$\sum_{1}^{2} (y_r + e_r)$
x_2			

Thus the cumulative relapse rate during $\underline{m}$ months as a percentage is:

$$\frac{100 \sum_{1}^{m} (y_r + e_r)}{x_o}$$

This should be equal to:

$$100 \left(1 - \prod_{1}^{m} (1 - q_r) \right)$$

in which $\underline{(1 - q_r)} = \underline{s_r}$ of the preceding development.

5. PARTICULARS OF TREATMENTS.

Cases which received Penicillin fall into the following groups:-

1. 2.4 mega units of Sodium Penicillin by three-hourly intramuscular injections of 40,000 Oxford units over a period of $7\frac{1}{2}$ days with no treatment other than local applications of saline or occasionally Eusol and other antiseptics to lesions;

2. 4.0 mega units of Sodium Penicillin given as above but for $12\frac{1}{2}$ days;

3. 2.4 mega units of Sodium Penicillin in $7\frac{1}{2}$ days with a course of intensive arsenotherapy beginning on the second day of Penicillin treatment, viz. Mapharside 0.06 gm. intravenously daily for a period of ten days;

4. 2.4 mega units of Penicillin, 0.4 gm. Mapharside and 1.0 gm. Bismuth given concurrently as above (3) over a period of 10 days.

Standard commercial preparations were used. The proportions of the different crystalline Penicillins in these preparations are not known. 'Long Term' Arsenic cases call for more detailed specification. Ostensibly, it was the policy to treat all cases of: (a) Primary Syphilis with at least 4 courses of Arsenic and Bismuth given concurrently in a twelve-month period, each course consisting of 5.85 gm. of N.A.B. and 2.0 gm. of Bismuth; (b) Secondary Syphilis

with at least 5 courses of Arsenic and Bismuth. In practice, it rarely proved possible to implement this procedure consistently. In fact, cases in our series treated with apparent success received from between 3-27 gm. N.A.B. and 3-18 gm. of Bismuth over periods varying from 8-23 months. Documentation w.r.t. some cases was poor, only the number of courses being stated. Other long-term cases received in addition to N.A.B. and Bismuth, two weekly injections of 0.06 gm. Mapharside in place of a single injection of 0.6 gm. N.A.B. Cases receiving Mercurials (3% of all) were rejected from the pool of data. The composition of the group was as follows:-

		% (437 cases)
N.A.B. Bi.	No. of courses and dosage known	61
N.A.B. Bi. Mapharside	No. of courses and dosage known	16.6
N.A.B. Bi. Mapharside	No. of courses known. Dosage not stated	22.4
		100.0

Cases which received Mapharside and cases for which treatment dosage was not fully stated make up more than a third of the total material available. So it was decided to retain

them. In specifying the equivalent dosage of N.A.B. received by such cases, we equate 0.12 gm. of Mapharside to 0.6 gm. N.A.B. and add the quantity of N.A.B. so computed to the actual amount of N.A.B. received.

TABLE III

Mean value in grams for N.A.B. and Bismuth and mean value in months for period of Treatment in cases receiving four courses of N.A.B. and Bismuth.

Diagnosis	No.of Cases	Mean dose N.A.B. (gm)	Mean dose Bismuth (gm)	Mean duration of Treatment (Months)
Prim: Sero: Neg:	36	19.5 ± 0.6	9.3 ± 1.3	13.2 ± 0.5
Prim: Sero: Pos:	66	19.0 ± 0.5	9.5 ± 0.4	13.7 ± 0.6
Secondary:	38	19.0 ± 0.2	9.7 ± 0.5	14.1 ± 0.5

From analysis (Table III) of dosage given in four courses to cases with reference to which the information was available, it was possible to distribute the L.T.A. cases among 2 groups:

(i) Group A cases receiving more than 19.0gm. of N.A.B. or, where dosage was not stated, four or more courses of Arsenic and Bismuth;

(ii) Group B cases receiving either less than 19.0 gm. of Arsenic or less than four courses of Arsenic and Bismuth. Tables IV A and B show that the only obviously significant difference between these groups is the dosage of arsenic:-

TABLE IV A

Analysis of cases receiving more than 19.0 gm. N.A.B. plus Bismuth.

Diagnosis	No. of Cases	Mean dose N.A.B. (gm)	Mean dose Bismuth (gm)	Mean duration of treatment (Months)
Prim:Sero: Neg:	28	22.1 ± 0.9	9.5 ± 0.6	14.1 ± 0.9
Prim:Sero: Pos:	50	22.1 ± 0.4	9.7 ± 0.2	13.7 ± 0.4
Secondary	52	22.3 ± 0.3	10.2 ± 0.4	14.7 ± 0.4

TABLE IV B

Analysis of cases receiving less than 19.0 gm. N.A.B. plus Bismuth.

Diagnosis	No. of Cases	Mean dose N.A.B. (gm)	Mean dose Bismuth (gm)	Mean duration of Treatment (Months)
Prim:Sero: Neg:	75	13.0 ± 0.5	8.7 ± 0.2	12.2 ± 0.4
Prim:Sero: Pos:	81	14.4 ± 0.3	9.8 ± 0.2	14.1 ± 0.4
Secondary:	53	13.4 ± 0.6	9.5 ± 0.4	14.0 ± 0.4

As a result of this analysis 79 cases receiving between 4 and 6 courses of Arsenic and Bismuth were added to Group A whilst 19 cases receiving between 2 and 3 courses only were added to Group B. No significant departures from these mean values occurred among cases which were known to have relapsed in the two groups. Analysis of Group A and B cases with reference to the incidence of jaundice revealed a highly significant difference (Table V) between the two groups. This was, of course, the prime reason for the lower average Arsenic dose in Group B. The incidence of dermatitis was also higher in this Group though it was a minor contributory factor in lowering the Arsenic dosage for the group. Differences with reference to incidence of jaundice between the three subdivisions within each group are not statistically significant, though examination of a larger series of cases has shown that maximum susceptibility to jaundice occurs during the very early stage of Syphilis.

TABLE V. % JAUNDICED

	No. of Cases	Primary Sero-Neg:	Primary Sero-Pos:	Secondary
Group A	209	15.2	27.8	20.5
Group B	228	54.3	60.6	45.2

A group of 'Long-Term' Arsenic cases subsequently treated with Penicillin started treatment on a purely 'Long-Term' Arsenic basis, but after a variable interval arsenotherapy was stopped; and they were given a routine course of Penicillin 2.4 Mega units in $7\frac{1}{2}$ days. Table VI shows the mean values for Arsenic and Bismuth received prior to Penicillin therapy and the mean period over which therapy (including Penicillin) took place.

TABLE VI

Mean quantities of Arsenic and Bismuth received prior to Penicillin Therapy.

Diagnosis	No. of Cases	Mean dose (N.A.B. (gm)	Mean dose (Bismuth (gm)	Mean duration of treatment (including Penicillin) (Months)
Prim:Sero: Neg:	26	7.7 ± 0.6	4.1 ± 0.4	6.4 ± 0.8
Prim:Sero: Pos:	39	7.1 ± 0.5	4.1 ± 0.5	7.0 ± 0.7
Secondary:	26	7.9 ± 0.7	4.8 ± 0.6	7.1 ± 0.1

It can be seen that similar amounts of N.A.B. and Bismuth were given throughout and that the period between the beginning of arsenotherapy and the end of Penicillin injections was not unduly variable. Amongst cases which are known to have relapsed there

were no significant departures from the mean values given above. The three divisions into Primary, Sero-negative, Primary Sero-positive and Secondary cases refers to the condition at the inception of Arsenotherapy. By the time Penicillin therapy began most of the initially Sero-positive cases had become Sero-negative. Table VII indicates the proportion of cases in which Sero-reversal had taken place, those in which the serum test was still positive and those who had not had a recent test.

TABLE VII. Serology at start of PEN: Therapy

Initial Diagnosis	No. of Cases.	NEG: %	POS: %	Insufficient Data %
Primary Sero-Neg:	25	100	0.0	0.0
Primary Sero-Pos:	38	64.1	10.3	25.6
Secondary	26	80.8	11.5	7.7

Cases which received Short-Term Arsenic Therapy fell into two groups. The first group comprised cases treated by daily intravenous injections of Mapharside of 0.06 gm. and intramuscular injection of Bismuth over a period of 20 days. On inspection the

cases consistently received a total of 2.0 gm. Bismuth and the treatment period was 20 days with exceptional and minor variations (e.g. ± 0.2 gm. Bismuth and ± 1 day respectively). Mapharside dosage varied between 1.2 gm. and 1.6 gm. for the course. Table VIII shows however, that the mean dosage was 1.3 gm:-

TABLE VIII.

Diagnosis	No. of Cases	Mean dose Mapharside (gm)
Primary Sero: Neg:	32	1.31
Primary Sero: Pos:	38	1.30
Secondary	42	1.29

The second group comprised only about 40 cases treated by semi-intensive courses of Mapharside, or Neohalarsen, and Bismuth over a minimum period of 30 days. The variability of dosage and treatment period were, however, too great to allow of standardisation in such a small group of cases. We therefore present the relapse rates for the 20-day treatment group only.

6. RESULTS.

Appendix 1 cites crude figures with reference to relapses, and failures and doubtful cases; and Appendix 2 cites cases remaining at the end of specified intervals after end of treatment. Table IX shows cumulative relapse rates at the end of 6 and 12 months for the different treatments. With regard to some, it was possible to follow sufficient cases for 18 months. Figures can be regarded only as trends owing to the small size of some samples. All failures as defined above have been excluded. It is evident that the Penicillin courses are comparatively unsuccessful except when following Long-Term Arsenic. The two L.T.A. schemes are apparently the most efficient particularly for treatment of Primary Syphilis and in contrast to other forms of therapy are the most successful treatments of Secondary Syphicis. Short-term Arsenotherapy takes an intermediate position. It is interesting to note the comparative relapse rates of S.T.A. (Mapharside 1.3 gm. and Bismuth 2.0 gm) with: (a) Penicillin 2.4 Mu + Mapharside 0.6 gm., and (b) Penicillin 2.4 Mu + Mapharside 0.4 gm. + Bismuth 1.0 gm.

Duration of treatment was much the same for each of these therapies. The relapse rates of S.T.A. are consistently smaller than those of either form of combined therapy. The least differ-

ence occurred in cases of Seropositive Primary and in cases of Secondary Syphilis between S.T.A. and Pen: 2.4 Mu and Mapharside 0.6 gm. Since Penicillin alone, in the two dosage schemes investigated, is least efficacious of all, these relapse rates suggest that Penicillin may in some way mitigate the effect of Arsenic on Treponema pallidum. It is however, necessary to point out that the dosage of Mapharside received by the S.T.A. group was at least twice as great as that given to the Penicillin cases.

Our data sustain the accepted conclusion that the prognosis for arsenotherapy is less favourable if treatment does not start until the sero-positive primary stage. Contrary to results reported by Moore and others they would also suggest that the prognosis for sero-positive primary cases is better than that for secondary cases. Our results do not suggest any consistent differences in these respects with regard to Penicillin therapy. Although attenuation of our material owing to the introduction of Penicillin shortly before demobilisation compels us to state only guarded conclusions with respect to the long-term prognosis of Penicillin therapy, the picture disclosed by Figure 1 suggests that the risk of relapse among cases treated in the primary stage of each is a limiting value of about 14% within the first twelve months.

TABLE IX.

PERCENTAGE RELAPSE RATES.

The sign ∓ indicates the variation from the mean of maximum and minimum values.

	Time Period	Pen 2.4 Mu	Pen 4.0 Mu	Pen 2.4 Mu and Maph. 0.6 gm.	Pen 2.4 Mu Maph. 0.4 gm. Bi 1.0 gm.	L.T.A.70 gm. Pen 2.4 Mu	L.T.A. (N.A.B.22 gm (A)	L.T.A. (N.A.B.14 gm. (B)	S.T.A. (Maph.1.3gm Bi 2.0 gm)
Primary Sero-negative	6 months	11.0 ∓ 0.5	(10.8 ∓ 0.8)	14.6 ∓ 1.5	5.0 ∓ 0.0	(0.0)	0.0	0.0	0.0
	12 months	14.0 ∓ 2.4	-	(18.2 ∓ 1.5)	(9.0 ∓ 1.9)	(5.7 ∓ 0.0)	(0.0)	(0.0)	(0.0)
	18 months	-	-	-	-	-	-	(0.0)	-
Primary Sero-Positive	6 months	10.2 ∓ 0.0	(6.3 ∓ 0.0)	6.7 ∓ 0.9	10.9 ∓ 0.0	2.6 ∓ 0.0	0.0	0.6 ∓ 0.6	5.3 ∓ 0.0
	12 months	12.8 ∓ 0.0	-	(15.9 ∓ 4.5)	(23.3 ∓ 3.3)	(2.6 ∓ 0.0)	1.6 ∓ 0.0	3.5 ∓ 0.6	(5.3 ∓ 0.0)
	18 months	(14.1 ∓ 1.3)	-	-	-	-	-	-	-
Secondary	6 months	8.1 ∓ 0.5	(17.4 ∓ 0.0)	(6.6 ∓ 0.0)	(19.2 ∓ 0.0)	(8.8 ∓ 0.0)	4.1 ∓ 0.9	3.8 ∓ 1.2	5.7 ∓ 0.0
	12 months	17.5 ∓ 0.5	-	-	-	(12.9 ∓ 4.1)	(11.4 ∓ 2.1)	(9.3 ∓ 2.8)	(10.7 ∓ 1.5)
	18 months	-	-	-	-	-	-	-	(16.8 ∓ 3.7)

Bracketed figures are based on sample of less than 30 cases.

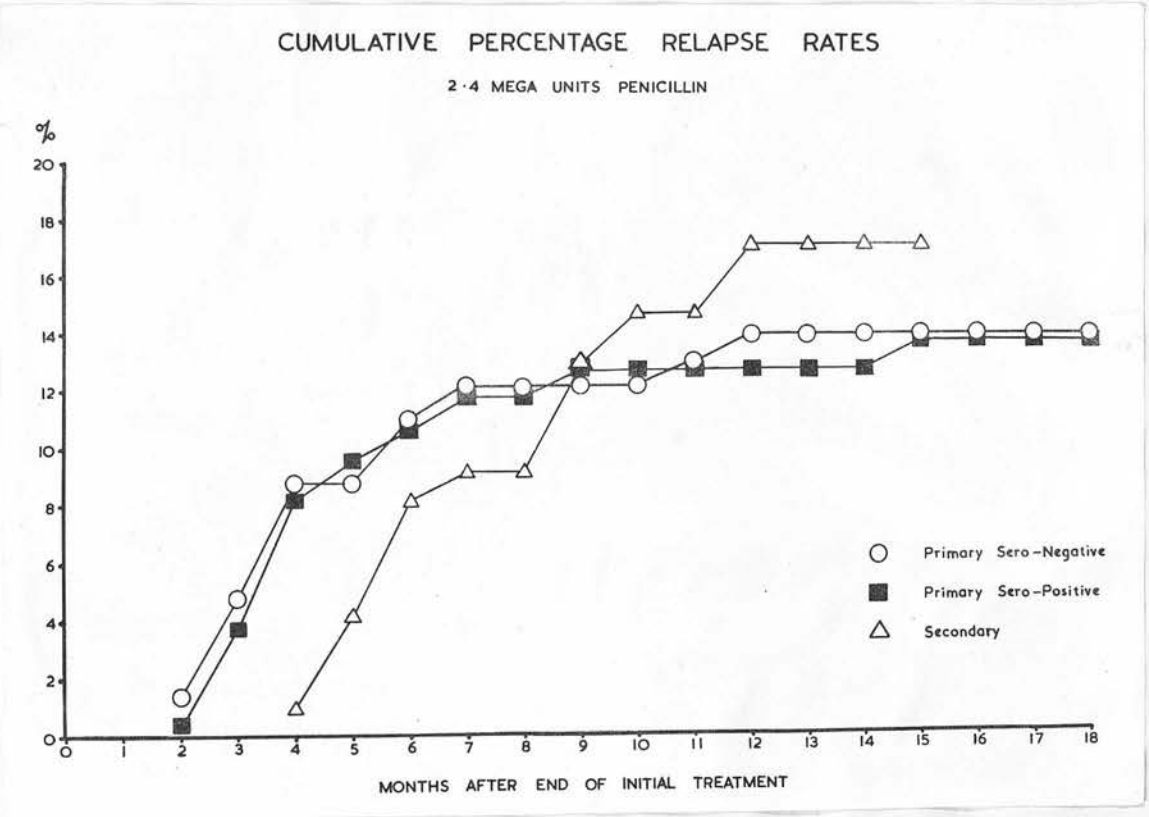
CUMULATIVE PERCENTAGE RELAPSE RATES
2·4 MEGA UNITS PENICILLIN
%
20
18
16
14
12
10
8
6
4
2
0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
MONTHS AFTER END OF INITIAL TREATMENT
Primary Sero-Negative
Primary Sero-Positive
Secondary

Fig.1

Our definition excludes the occurrence of failures among Sero-negative Primary cases. Table X shows an analysis of those cases remaining Sero-positive for four or more months. Penicillin has a comparatively higher rate of failure among Sero-positive Primary cases but among Secondary cases, particularly when combined with Mapharside, compares not unfavourably with S.T.A. and the smaller dosage group (B) of L.T.A. For most treatments the prognosis in this respect of secondary cases is less favourable than the prognosis of primary cases.

TABLE X. ANALYSIS OF TREATMENT FAILURE CASES.

Percentage of treatment failure cases (defined as cases remaining persistently sero-positive for four months or more after treatment) as a percentage of total numbers of cases (within each treatment).

	PEN: 2.4Mu	PEN: 4.0Mu	PEN:2.4 Mu and Maph: 6.0	PEN:2.4 Mu Maph: 0.4 gm. Bi 1.0 gm.	LTA 7gm & PEN: 2.4 Mu	LTA (A) 22 gm	LTA (B) 14 gm	STA 1.3 gm. Bi 2.0 gm.
Primary Sero-Positive	2.9∓0.0	1.6∓1.6	0.8∓0.8	5.5∓0.8	0.0	0.0	0.0	0.0
Secondary	7.8∓3.2	3.8∓3.8	2.5∓2.5	2.8∓2.8	0.0	0.0	5.0∓0.0	10.7∓6.0

SUMMARY.

Relapse rates during a 12-month period of cases of Early Syphilis respectively treated with eight different therapies are contrasted and disclose the following conclusions:

1. Long-term Arsenic in a mean dosage of 22 gm. N.A.B. given in four or more courses proved to be the most effective therapy for all forms of Early Syphilis. In dosages averaging 14 gm. in fewer than four courses, it was the next most efficacious.

2. For Primary cases in the doses, and for the period, specified in the test Penicillin alone was inferior to either long-term or short-term Arsenotherapy both with respect to prognosis of relapse and resistance to treatment.

3. The data of this enquiry give no conclusive indications of synergism with reference to simultaneous use of Penicillin and Arsenicals with Bismuth, being suggestive, if at all, of the reverse.

4. Though the data recorded admittedly antedate recognition of the importance of the K-fraction, the results at least emphasize the need for long-term follow-up before replacing arsenotherapy by Penicillin.

5. On the other hand, it is necessary to admit the possi-

bility that greater frequency and/or duration of Penicillin treatment in contradistinction to greater aggregate dosage might reverse the unfavourable verdict which our data compel us to record.

REFERENCES:

HUBELO et RABUT:	La Presse Médicale, 1924.
MOORE, J. E.:	The Modern Treatment of Syphilis, 1943.
TRUELOVE, S.C. and HOGBEN, L.:	J. Soc. Med. 1947, I, 1.

APPENDIX 1.

The relative distribution of relapses and treatment failures in the various treatment groups and their subdivision into primary sero-negative, primary sero-positive and secondary cases. Space does not permit of indication of the relationships of relapses to time intervals after treatment. Crude figures.

	PRIMARY SERO-NEGATIVE			PRIMARY SERO-POSITIVE					SECONDARY				
	Clin. re-lapse	Definite Sero-relapse	Doubtful Sero-relapse	Clin. Re-lapse	Definite Sero-relapse	Doubtful Sero-relapse	Definite Treatment failures	Doubtful Treatment failures	Clin. Re-lapse	Definite Sero-relapse	Doubtful Sero-relapse	Definite Treatment failures	Doubtful Treatment failures
PEN: 4	11	1	3	15	3	1	4	0	8	4	2	5	7
PEN: 4.0	5	0	1	1	0	0	0	1	2	1	0	0	1
PEN: & MAPH:	10	1	2	5	1	. 3	0	1	3	0	0	0	2
PEN:MAPH: & Bi.	5	0	1	6	1	1	3	1	6	1	0	0	2
LTA & PEN:	2	0	0	2	0	0	0	0	0	3	1	0	0
LTA (A)	0	0	0	0	1	0	0	0	0	0	6	0	0
LTA (B)	0	0	0	1	1	0	0	0	0	4	0	2	0
STA	0	0	0	2	2	2	0	0	1	5	0	2	5
TOTAL:	33	2	7	32	9	7	7	3	20	18	9	9	17

APPENDIX 2.

Numbers of cases in each treatment group (sub-divided into primary sero-negative, primary sero-positive and secondary) and the numbers of cases remaining under surveillance after six and twelve months.

It will be noted that relatively small numbers of cases in the LTA, LTA + PEN and STA groups were, in general, followed over a longer period than the other cases.

		PEN: 2.4	PEN: 4.0	PEN: MAPH:	PEN: MAPH: & Bi.	LTA & PEN:	LTA (A)	LTA (B)	STA	
Primary Sero-Negative	Total treated	114	57	108	105	25	44	75	32	
	Remaining 6/12	80	20	30	51	23	30	64	31	
	Remaining 12/12	41	0	6	5	8	10	34	23	
Primary Sero-Positive	Total treated	163	41	110	89	38	84	84	38	
	Remaining 6/12	122	14	31	42	31	64	63	35	
	Remaining 12/12	52	0	6	2	17	31	31	29	
Secondary	Total treated	118	28	53	48	26	67	46	42	
	Remaining 6/12	86	8	25	20	19	48	36	39	
	Remaining 12/12	32	0	5	4	9	26	20	28	TOTAL CASES
TOTAL TREATED		395	126	271	242	89	195	205	112	1635

PART III. THE USE OF PALUDRINE FOR B.T.MALARIA RELAPSES.

An enquiry to compare Mepacrine therapy and Quinine and Pamaquin therapy was reported in the Bulletin of Army Health Statistics (1st Series 11th Issue) in September 1945. By arrangement with Consulting Physician a further comparative trial was arranged to test Paludrine. The writer took responsibility for the analysis of the data and design of the records. The clinical details of diagnosis and treatment were carried our by Major R.D.C. Johnstone, R.A.M.C. Medical Specialist at Colchester Military Hospital, who also dealt with follow-up of cases during the six months after discharge from Hospital. The three methods of treatment under consideration were:

A. Quinine gr. x and Pamaquin 100 mgm. three times a day for 10 days.

B. Paludrine (M.4888) 25 mgm. twice daily for 10 days.

C. Paludrine (M.4888) 250 mgm. twice daily for 10 days.

All cases in the series were male patients admitted to the Colchester Military Hospital in the normal routine and were unselected except in so far as they were proved to have B.T. Malaria by positive blood smear. Each treatment was administered to every third patient in order of diagnosis; and no man was selected for

a particular treatment because the physician considered that more suitable for him. All were treated alike in the ward; and no discrimination was permitted to allow patients to know what drugs they were being given.

The previous enquiry confirmed an already recognised periodicity of relapse with a major peak in the second and a minor peak in the ninth month. To establish a high order of therapeutic efficacy for a new treatment, it would therefore be necessary to follow the history of a group of patients through a period of at least one year. On the other hand, the validation of a negative finding needs no such exacting test. That relapses occur in excessive proportions during the early months after treatment suffices to dispose of the claim that a treatment is highly efficacious. It happens that observations on the results of Paludrine administration during the first six months satisfy this criterion. To provide an answer to the question which prompted this enquiry it has not therefore been necessary to await results of a more protracted trial. Follow-up was by written questionnaire and where ambiguous replies were given, a further letter was sent to elucidate the query. Of 322 cases who received it, satisfactory information was received from all but 5 cases. Information with reference to the others left no reason to doubt whether there had been a recurrence of the disease or whether the patient had taken therapeutic or suppressive

medications liable or likely to postpone relapse.

These 322 cases were thus distributed with respect to treatment:

A.	107	-	including 3 doubtful.
B.	108	-	" 1 "
C.	107	-	" 1 "
	322		

Analysis of the results 6 months after the end of treatment (Table I) showed that the Relapse Rate following the standard treatment with Quinine and Pamaquin was significantly less than the relapse rate for either of the two groups receiving Paludrine. There was no significant difference with respect to relapse rates of patients respectively given large or small doses of Paludrine over a period of ten days.

TABLE I.

RELAPSE RATES 6 MONTHS AFTER TREATMENT

Scheme	Total Cases	% Doubtful	% Relapses	Differences
A	107	2.8	19.6 ± 3.8	25.8 ± 6.1
B	108	0.9	45.4 ± 4.8	29.0 ± 6.1
				3.2 ± 6.8
C	107	0.9	48.6 ± 4.8	

The relapse rate for Quinine and Pamaquin in Table I compares unfavourably with the rate of 10.3% after five months in the earlier series referred to above; and therefore calls for comment. The extra month of the present follow-up can account for few additional cases; and the difference is at least in part explicable, if we pay due regard to the different methods of follow-up employed in the two enquiries. In the present series the personal questionnaire by the physician evoked a definite reply from each patient. The method of collecting material for the earlier enquiry gave a reliable comparison between the two groups of cases within the series; but provided no firm basis for absolute rates. Hence the results are not comparable with those of any other series.

With a view to assessing the effect of treatment given early or late in relation to onset of symptoms, the mean duration in days from onset of symptoms to start of treatment was determined for each group. That of Group A was somewhat less than that of B or C, the difference being significant. Among those who received Quinine and Pamaquin, a relatively high proportion therefore received it (Table II) at an early stage after the onset of the attack. This might conceivably introduce a bias into comparison of figures shown in Table I. Accordingly Table II shows relapse rates for different time intervals between onset of symptoms and start of treatment. It omits cases treated during the first 24 hours (one case in Group

A, two in B, one in C) and certain cases where the time intervals were uncertain. In conformity with the results of a more refined breakdown of relapse rates with reference to the time interval under discussion the grosser comparison which Table II exhibits demonstrates that Group A consistently had lower relapse rates than Groups B or C treated during a comparable period after onset of symptoms.

TABLE II

RELATION OF THE TIME INTERVALS BETWEEN ONSET OF SYMPTOMS AND STARTING TREATMENT.

Duration	% CASES			% RELAPSES		
	Scheme A (102)	Scheme B (101)	Scheme C (102)	Scheme A	Scheme B	Scheme C
24-48 hours	12.7	6.9	6.9	23.1	28.6	71.4
49-96 hours	51.0	39.6	34.3	19.2	50.0	54.3
97-144 hrs.	24.5	33.7	35.3	16.0	50.0	41.7
Over 144	11.8	19.8	23.5	16.7	35.0	45.8
	100.0	100.0	100.0			

In so far as the figures of Table II suggest that there is any relation between therapeutic efficacy and what time elapses from onset of the attack to start of treatment those specifically referable to Group A might encourage the suspicion that too early treatment favours

a high relapse rate. If so, the higher proportion of Group A patients treated during the first interval would bias the result of an unduly low estimate of the relative efficacy of Quinine and Pamaquin. An alternative analysis of the same data as shown below reinforces this suggestion. Though the differences are not statistically significant, they point to an issue worthy of further enquiry and of consideration in the design of any future therapeutic trial on anti-malarial drugs.

25 - 48 hours	23.1% ± 12.2
49 - 72 hours	21.1% ± 9.6
73 - 96 hours	18.2% ± 6.8
97 -120 hours	9.5% ± 6.8

One other possible source of bias calls for comment, i.e. the relation of numbers of previous attacks to the efficacy of any particular treatment. Table III discloses no indication that cases with a history of four or more previous relapses are more likely to relapse after treatment than cases with a history of fewer than four previous relapses.

TABLE III
(Quinine and Pamaquin Cases)
RELATION OF RELAPSE RATES TO NUMBERS OF PREVIOUS ATTACKS OF B.T. MALARIA.

	No. of Cases	% Relapse	Difference
3 or fewer attacks	29	17.2	
			4.2 ± 8.9
4 or more attacks	75	21.3	

CONCLUSIONS:

1. During the first six months after treatment between 45% of B.T. Malaria relapses relapsed again after treatment with 25 mg. of Paludrine, twice daily for 10 days and 49% similarly relapsed when treated with 250 mg. doses.

2. The corresponding relapse rate for a control group treated with Quinine and Pamaquin was less than 20%.

3. In any future therapeutic trial of anti-malarial drugs, it might be profitable to direct attention to the possible relation between the efficacy of a treatment and the interval to its initiation from the onset of symptoms.

PART IV. TREATMENT OF INTERNAL DERANGEMENT OF THE KNEE.

1. INTRODUCTION.

Among men in the Army I.D.K. ranks as a major source of wastage alike with respect to incidence, man-days lost to service and invalidings. From the figures available for 1943 it appears that the absolute incidence in Britain was about 0.3%. The relative morbidity rate was 1.7% and the relative wastage rate 4.1%. That is to say I.D.K. accounted for over 4% of man-days lost to service on account of sickness. In the same year it accounted for 0.5% of discharges on account of sickness from the Army as a whole. Clearly, it is therefore important to assess circumstances contributory to its occurrence and what measure of success attends its treatment. Such is the aim of what follows.

The ensuing analysis refers to a nominal roll of Military O.Rs. admitted to hospital in Britain during the 12-month period August 1943 - July 1944. The size of the sample was 777 of whom 50 were admitted twice within this period. During their Army service up to July 31st 1944, some of these 777 cases had been previously hospitalised for the same complaint, and as the figures below show, the 777 individuals studied thus correspond to a grand total of 967 admissions to hospital:

Admissions per Individual	Individuals	Man-admissions
1	627	627
2	118	236
3	25	75
4	6	24
5	1	5
Total:-	777	967

2. DURATION OF TREATMENT IN DIFFERENT MEDICAL UNITS.

Two striking features of wastage with respect to I.D.K. exposed in the ensuing table are:

(a) the protracted period between hospital discharge and return to unit of cases whether treated by operation or not;

(b) the more protracted stay of cases in E.M.S. as compared with Military hospital.

As regards (b) it is noticeable that the longer duration of hospital treatment among patients who underwent operation in E.M.S. Hospitals is not associated with longer convalescence.

Analysis of Days away from Unit, Mean days per man

	WITHOUT OPERATION (35.4%)		WITH OPERATION (64.6%)	
	E.M.S.	Military Hospital	E.M.S.	Military Hospital
Hospital	31.2	18.4	61.6	42.9
Con. Home	28.9	32.7	39.7	42.1
Con.Depot	43.7	40.9	51.2	51.1
Total Duration	54.7	29.7	120.5	104.8

Owing to incompleteness of information and the fact that all patients do not necessarily go through Convalescent Homes and Depots, these figures are calculated from all records with reliable data, and the individual items in each series are not necessarily based on the same cases or the same number of cases. Comparison of the two types of hospital discloses a discrepancy with respect to length of stay in hospital and total duration of absence from duty of men treated without operation. As regards the former, the difference between patients in E.M.S. and patients in Military Hospitals is 12.8 days. As regards the latter the difference is 25 days. Duration of stay in Convalescent Homes and Depots is approximately equal for both classes of patients. The discrepancy is due to transference of fewer cases from Military Hospitals to the Y-list and hence fewer to Convalescent Depots. This is demonstrable by contrasting: (a) mean duration of stay in the two types of hospital (Military 18.4 days, E.M.S. 31.2 days); (b) the proportion of patients respectively sent from each type of hospital to Convalescent Depots (Military 13.8%, E.M.S. 42.3%). Longer stay in E.M.S. Hospitals might also be attributable in part to delay of operation or longer conservative treatment in E.M.S. as compared with Military Hospitals, of which a greater proportion have out-patient departments for preliminary examination and surveillance. The figures shown do not conspicuously confirm this supposition with reference to cases subjected to operation.

Days in Hospital before operation	MEAN DAYS PER CASE	
	E.M.S.(286)	Mil.Hospital (176)
21 or less	7.1 (237)	6.3 (148)
Over 21	48.4 (49)	39.2 (28)

3. CRITERIA OF EFFICACY.

Of statistical indications with respect to success of treatment three call for special comment: (a) indications of previous operation; (b) discharge from the service on medical grounds; (c) change of medical category. As regards (a) it is noteworthy that only 14 out of 500 menisci removed were remnants (11 posterior horns and 3 rims) of previous meniscectomies. Though I.D.K. is a major source of man-days lost to service it does not in fact make a considerable contribution to total wastage, i.e. discharge under category E, and such was the disposal of only 1.8% of the sample dealt with. Change of medical category (if stated) of cases treated or not treated surgically might thus be regarded as the most instructive indication of its efficacy at our disposal; such information is available with respect to the patient at the time of admission to hospital and return to unit after discharge from hospital or Convalescent Depot. The following table refers to cases with respect to which we have information concerning medical category both at admission and on discharge.

	WITH OPERATIVE TREATMENT		WITHOUT OPERATIVE TREATMENT	
	No. of Cases	% without lowering of medical category	No. of Cases	% without lowering of medical category
E.M.S.	32	53.1 ± 8.8	77	52.0 ± 5.7
Military) Hospital)	30	70.0 ± 8.4	29	44.8 ± 9.2
All	62	61.3 ± 6.2	106	50.0 ± 4.9

None of the differences recorded in this table is significant. Accordingly, there is no indication that conservative treatment has a bad prognosis, but this may merely be attributable to the fact that cases not recommended for surgical treatment were the less severe ones.

PART V. SULPHA DRUGS AND PENICILLIN FOR THE TREATMENT OF GONORRHOEA.

1. DOCUMENTARY SOURCES.

What follows refers to cases (Male Other Ranks) of Gonorrhoea treated in Hospitals in Britain between January 1943 and September 1945. Information was available with respect to diagnosis, age, duration of stay in hospital, courses of treatment by date, presence or absence of relapse and interval from completion of stay in hospital to day of subsequent admission, where relevant, and previous history with respect to the disease.

2. CRITERA OF THERAPEUTIC EFFICACY.

Three criteria are practicable:

(a) response to a single course of treatment:

(b) relapse rate:

(c) duration of stay in medical unit.

(a) Response to a single course of treatment. A common difficulty of enquiries of this sort arises from the fact that the physician's primary concern is to cure the patient by any means at his disposal rather than conform to the requirements of a controlled experiment. Hence many patients receive more than one method of

treatment, concurrently or successively. For this reason it has been convenient to classify the records with respect to treatment in two major groups:

(A) those who received initially a single course of:-

(a) penicillin, (b) sulphathiazole, (c) sulphapyridine;

(B) those who received treatment with one or other of the above and simultaneous irrigation with either $KMnO_4$ or HgOCN.

The operative word in (A) is initially, because any case which fails to respond to a single course of therapy normally receives additional treatment. This circumstance permits us to refine our criterion of success. Since no case is ordinarily discharged from hospital without a clean bill, the use of a single method or simultaneous combination of two methods followed by no other is a sufficient indication that such a course is successful. Conversely, recourse to additional courses after a lapse of time sufficient to assess the value of the primary treatment is one admissible criterion of failure.

From this it follows that our data provide no clear guidance with respect to the efficacy of a combination of successive, as opposed to simultaneous therapies. Individuals who receive a subsequent course of treatment may be highly selected, both because they are patients who fail to respond to a single method and because the Venereologist will exercise his clinical judgment to prescribe the

choice of further treatment.

(b) Relapse Rate. As a criterion of efficacy on its own merits the relapse rate, here defined as readmission within 60 days, calls for no comment; but it is possible to apply a more exacting yardstick of cure than the one specified in preceding remarks by combining information both with respect to failure to respond to a single (unitary or composite) course of treatment, and to recurrence after discharge from hospital. What we here designate a permanent cure therefore signifies a case: (a) completed without recourse to subsequent treatment: (b) not admitted again to hospital within the arbitrary period of 2 months, delimited with due regard to the danger of confusing a genuine relapse with a reinfection.

(c) Duration of stay in Medical Unit. With due regard to variation arising from administrative instructions, duration of stay may be a useful yardstick of therapeutic efficacy if a condition involves protracted hospitalisation; but the inevitable delays involved in transfers of patients to and from medical units exclude any reliance upon it, when the total time spent in hospital is about a week. Hence we should not expect any conspicuous difference between treatments to show up in the hospital sample tabulated below.

TABLE II.

DURATION OF STAY IN HOSPITAL - SIMPLE TREATMENTS.

	No. of Cases	Mean Days	Differences	
PENICILLIN	215	6.8 ± 0.8		
			2.0 ± 1.2	
SULPHATHIAZOLE	358	8.8 ± 0.9		3.0 ± 0.8
			1.0 ± 0.9	
SULPHAPYRIDINE	113	9.8 ± 0.2		

Data from another source records interval between admissions to forward medical units and cessation of urethral discharge. These disclose a striking difference, and hence show how unavoidable delays inherent in hospital administration attentuate such figures as those in Table II. The relevant ratios exhibited below show that simple treatment with penicillin excels simple treatment with sulpha drugs with respect to the rapidity of the cure as well as the proportion of successes.

TABLE 12.

DURATION OF TREATMENT IN FORWARD UNITS.

	No. of Cases	Mean Days	Differences
PENICILLIN (alone)	598	1.9 ± 0.004	
			3.9 ± 0.006
SULPHA DRUGS (alone)	266	5.8 ± 0.017	

PART VI. PENICILLIN TREATMENT OF IMPETIGO.

The hospital records of 1031 soldiers suffering from Impetigo who were treated in Britain during 1943-44 were examined to compare different methods of treatment: of these methods the documents admit a threefold classification as follows:

(i) Over half (52.2%), here denoted Group O, received only lotions, ointments, compresses and drugs other than those of the sulpha group:

(ii) A group (S) of 425 cases (41.2%) received sulpha drugs, e.g. microcrystalline sulphathiazole and sulphanilamide powder:

(iii) A small residual group (P) of 68 (6.6%) received penicillin.

The clinical notes indicated that some of the cases had at some time previously suffered from the disease. Over 80% were apparently new. As shown below, the proportions of new cases in the three groups were not significantly different:

	Group O.	Group S.	Group P.
New Cases	80.7 ± 1.7	83.7 ± 1.8	79.4 ± 4.9
Relapses	19.3	16.3	20.6
TOTAL:	100.0 (538)	100.0 (425)	100.0 (68)

In the absence of more specific indications provided by clinical notes on the case record, Army medical documents furnish

the following yardsticks for the assessment of therapeutic procedures:

(a) Mortality rates;

(b) Relapse rates;

(c) Rate of subsequent discharge from the Service on medical grounds;

(d) Supervening diseases and/or complications;

(e) Duration of hospitalisation before being returned to Unit;

(f) Subsequent medical recategorisation.

It is rarely practicable to apply all of the above. If a disease is not a serious source of <u>total</u> wastage neither (a) nor (c) is practicable. Impetigo is not fatal. It is rarely a sufficient reason for discharge from the Army; and is not commonly associated with supervening disease or characteristic complications. Administrative recommendations to ensure reliable information respecting medical recategorisation have not yet taken effect. We are thus left with (b) and (e). Though figures cited above suggest that the relapse rate is not negligible, relapses which occur in the Army are trivial. Our only practicable yardstick is therefore (e). Mean duration of stay is admittedly a criterion liable to bias resulting from administrative considerations which govern hospital policy or disposal; but if the classes of a large sample

are contemporaneous, the association of shorter mean duration with a given treatment constitutes a strong prima facie case for its greater intrinsic efficacy. The following figures exhibit the mean duration (days) of hospitalisation for each of the 3 groups taken as a whole (i.e. including both new cases and cases with a previous history of the disease):

	Mean Duration of Stay.	Difference
Group O (neither sulpha drugs nor penicillin) ...	19.5 ± 0.68	
		3.8 ± 0.84
Group S (sulpha drugs) ...	15.7 ± 0.50	
		3.6 ± 0.92
Group P (penicillin) ...	12.1 ± 0.77	

Each of the differences recorded above is significant, i.e. (a) duration of stay with respect to penicillin treatment is significantly shorter than with respect to treatment with sulpha drugs; (b) duration of stay with respect to treatment with sulpha drugs is significantly shorter than with respect to treatment with neither sulpha drugs nor with penicillin. Though the differences as regards the composition of the sample, in so far as the proportion of putatively new cases might bias the result, are not statistically significant, the issue calls for a check. A similar analysis of new cases only yields the following results:

	Mean Duration of Stay.	Difference	No. of Cases
Group O ...	18.3 ± 0.7		434
		3.1 ± 0.86	
Group S ...	15.2 ± 0.5		356
		3.4 ± 0.86	
Group P ...	11.8 ± 0.7		54

A similar analysis of the chronic class yields the following:

	Mean Duration of Stay.	Difference	No. of Cases
Group O ...	24.5 ± 2.0		104
		6.4 ± 2.4	
Group S ...	18.1 ± 1.3		69
		4.7 ± 2.8	
Group P ...	13.4 ± 2.5		14

All the differences in the last two sets of figures tally, and all are statistically significant except the difference between the S and P relapses. The dubious significance of the latter difference is sufficiently explicable on account of the very small number of cases in the P group. However the difference with respect to mean duration of stay after penicillin treatment of the chronic cases (denoted as relapses above) and treatment of such individuals with neither penicillin nor sulpha drugs is highly signi-

ficant, being 11.1 ± 3.2.

As we might well expect, the preceding breakdown suggests that chronic cases are more recalcitrant to treatment. It is therefore fitting to exhibit the two sets of figures in a different way:

	New Cases	Chronic Cases	Difference
Group O ...	18.3 ± 0.7	24.5 ± 2.0	6.2 ± 2.1
Group S ...	15.2 ± 0.5	18.1 ± 1.3	2.9 ± 1.4
Group P ...	11.8 ± 0.7	13.4 ± 2.5	1.6 ± 2.6

Preceding remarks have drawn attention to one disadvantage of using duration of stay as the only practicable yardstick of therapeutic efficacy, in so far as it is liable to bias arising out of differences of administrative procedure. It is therefore important to take cognizance of the fact that some patients received treatment in E.M.S., others in military hospitals. This consideration is particularly pertinent, because the mean duration of stay for skin diseases is somewhat greater in the former than in the latter. Hence differential distribution with respect to the two groups of hospital in groups subjected to different methods of treatment might well account for the different results recorded above, if the proportionate contribution of E.M.S. hospitals was greater for treatment groups with longer duration of stay. Analysis of the sample from this point of view reveals that this is so. Fortunately, the diff-

iculty is easy to short-circuit because the majority of cases dealt with received treatment in military hospitals, and these groups alone suffice to provide a homogeneous sample large enough to act as a check on conclusions seemingly justified by figures exhibited in preceding tables. The figures below refer to fresh infections treated only in military hospitals:

Fresh Infection - Military Hospitals.

	Mean Duration of stay	Difference of means	No. of Cases
Group O ...	17.69 ± 0.73		312
		3.18 ± 0.93	
Group S ...	14.51 ± 0.57		277
		2.97 ± 0.94	
Group P ...	11.54 ± 0.74		48

A similar analysis of chronic cases treated in military hospitals yields the following results:

Chronic Infections - Military Hospitals.

Group O ...	24.5 ± 2.27		87
		5.9 ± 2.6	
Group S ...	18.6 ± 1.38		59
		5.6 ± 2.6	
Group P ...	13.0 ± 2.18		15

The foregoing analysis of cases treated only in military hospitals confirms all the conclusions which a corresponding treatment of the whole sample appears to justify; and thus removes any misgivings prompted by the inclusion of cases treated in hospitals of the two types. As before we find that the chronic cases are more recalcitrant:

	New Cases	Chronic	Difference
Group O ...	17.7 ± 0.73	24.5 ± 2.27	6.8 ± 2.4
Group S ...	14.5 ± 0.57	18.6 ± 1.38	4.1 ± 1.5
Group P ...	11.5 ± 0.74	13.0 ± 2.18	1.5 ± 2.3

In so far as we are entitled to regard mean duration of hospitalisation as a valid yardstick of therapeutic efficacy, the foregoing analysis clearly demonstrates the superiority of sulpha drugs to agents previously in use and the superiority of penicillin to both. From an administrative standpoint, we may approach the issue in a different way by exhibiting the figures to bring into clearer perspective the proportion of individuals returned to Unit within a given period, as below.

	ALL HOSPITALS			MILITARY HOSPITALS ONLY		
	10 days	20 days	30 days	10 days	20 days	30 days
Group O	30.0	73.4	84.5	31.8	68.2	84.5
Group S	34.3	82.8	87.4	35.7	83.9	91.7
Group P	47.1	94.1	97.0	49.2	93.7	96.8

PART VII. GENERAL CONCLUSIONS.

The remainder of this thesis will deal seriatim with general issues raised by the separate investigations recorded in Parts II - VI.

A. Syphilis Therapy in the Army. (Part II).

The calculation of Relapse Rates in this trial presented a major problem which is not uncommonly encountered in assessing results of Therapeutic Trials. The problem does not usually arise in ad hoc enquiries as defined above (P. 6); but the trial on Syphilis therapy was only in part such an enquiry. On the whole it must be regarded as a post hoc Enquiry. As already stated, the principal source of information was the Central Syphilis Register maintained by the Army Medical Directorate. This is a register of all men treated for Syphilis, and by means of it, serological surveillance is maintained for a period of two years from the end of treatment. The register contains notes of initial diagnosis and treatment together with results of subsequent serological tests which are performed at statutory intervals. A follow-up is thus possible of all patients whether they relapse, are re-infected or apparently cured. In general the system works well;

but failure in follow-up may arise from any of the following causes:-

1. Death from any cause other than Syphilis
2. Desertion from the Army
3. Clerical errors in reporting to the Central Register either by omission to forward documents or by giving incorrect information.
4. Default of the patient to report for surveillance regularly.
5. Loss of documents in transit.
6. Demobilisation of the patient.

Nothing can be done to rectify the first two items but clerical errors and defaulting can be corrected to some extent because the Central Register is constantly audited to ensure that patients do continue to have serological tests; and there is communication with the treatment centre concerned with the man's follow-up to obtain the missing information. Since treatment centres keep abstracts of treatment and serology of all patients on their register, loss of documents in transit may also be rectified. Loss of cases in a follow-up from any of the foregoing causes thus involves very small numbers; and there is no known selection to bias results if such incomplete cases are omitted from a survey. On the other hand the sixth of the above items constitutes a grave source of wasted cases at the present time, since demobilised men cannot be followed into civilian life. In other respects, the Syphilis

trial demonstrates the advantages of medical surveys among an army population. A moderate amount of pressure can be brought to bear on patients in a way which is impossible among civilians. Economic factors such as weigh heavily on the individual civilian patient, and shortage of beds as occurs in civilian hospital practice do not apply to the military patient who is not discharged from hospital until he is fit to return to duty.

In this survey of Army Syphilis Therapy carried out on patients treated by several different methods since 1943, it was possible to follow many cases for two years, but many more were lost owing to demobilisation. In the past, the results of Syphilis Therapy have been only roughly assessed, no long-term follow-up being possible. Since relapses of sero-negative patients after Arsenic Therapy may occur as much as two years subsequent to the end of treatment, possibly longer, loss of cases in a follow-up is most important. Such lost cases may be regarded in several ways:-

1. They all remain sero-negative
2. They all become sero-positive
3. They become sero-positive in the same proportion as the cases which have been successfully followed-up.

The last assumption is the most likely to give the best estimate; but if relapse is very late the numbers actually remaining

in the follow-up may be so few as to yield figures subject to condieerable sampling error. The adjustment of the relapse rate to take into account the speed of demobilisation presents a difficult problem comparable with the computation of a life table for a community subject to large-scale immigration; and is not, as already indicated, an insurmountable problem; but the taxonomical problem of delimiting relapses as such is beset with several pitfalls.

In determining criteria of Relapse it was found to be exceedingly difficult to distinguish relapse from re-infection. Histories are often untrustworthy and cannot be relied on. The exact site of the original sore may not have been noted; and it cannot be said that a new sore will not occur by re-infection on the site of an old one. Therefore, any case with evidence of renewed disease either by serological test only or by other clinical manifestations was classified as a relapse. This practice is reasonable and is usual among Syphilologists.

With respect to the choise of a Control Group, each of the different treatments provides a basis for comparison with each of the others. Only control of this sort is possible in a community which would not condone failure of the clinician to treat the patient in the interest of scientific curiosity. As an example of published work the longest follow-up is noted in a report from Johns Hopkins Hospital, Baltimore. It was a five-year follow-up

of 551 patients suffering from Early Syphilis (primary, early secondary and delayed secondary) most of whom were treated with Arsphenamine but of whom 17 received no treatment at all. After five years the cumulative relapse rate was 34.3%. This series was obviously not comparable, because: (a) it includes a wider range of cases, viz., delayed secondary syphilis; (b) the cases are not homogeneous; some were treated, some were not.

B. The Use of Paludrine for B.T. Malaria Relapses (Part III).

Two issues arising out of this trial call for additional comment:

1. A therapeutic trial must be regarded as a trial of the drug only in the specific conditions of the trial and in the dosage in which the drug is used. Large or small doses of Paludrine given twice daily for ten days only are to all intents and purposes ineffective as a cure of Benign Tertian Malaria; but we are not entitled to dismiss its usefulness without further consideration. Indeed, excellent results have since been obtained by other workers giving a small dose twice a week for 6-9 weeks.

2. Samples can be made homogeneous only by taking care with regard to those factors known to affect the experiment and those factors which, there is scientific reason to believe, might

affect the experiment. In this instance different phases of the life-cycle of the parasite go with different signs and symptoms. Since the parasite may well be more or less susceptible in different phases of the cycle, timing of the drug may well affect its efficacy. From the results of the trial, it did appear that the timing of therapy relative to the duration of symptoms was important and that Scheme A was biased by having a majority of cases treated early. The following table shows the mean interval from onset of symptoms to the beginning of treatment for each series of cases. There is a statistically significant difference between the mean durations for Schemes A and B and for Schemes A and C:-

		No. of Cases	Mean days	Differences and standard errors of the Mean
Schemes	A	103	3.4	0.8 ± 0.35
	B	103	4.2	
Schemes	A	103	3.4	1.1 ± 0.35
	C	103	4.5	
Schemes	B	103	4.2	0.3 ± 0.41
	C	103	4.5	

Table II of the Report shows that this source of bias did not suffice to distort the results of the comparison, since Group

A consistently had lower relapse rates than Groups B or C during a comparable period after onset of symptoms.

This illustration of unforeseen difficulties of obtaining a homogeneous sample in spite of precautions used in selecting comparable patients for the experiment shows the danger of relying on judgment in advance to guarantee homogeneity of small samples. The particular defect noted here was one which would almost certainly have been eliminated in a larger series of cases.

C. Treatment for Internal Derangement of the Knee (Part IV).

This survey of hospital records of men in the British Army admitted for treatment of Internal Derangement of the Knee was carried out partly as a therapeutic trial and partly to determine the pathology of the condition, its causes, and methods and effects of treatment. It indicates the weakness of Duration of Hospitalisation as a criterion of efficacious treatment when hospitals involved have different administrations. It also shows that medical re-categorisation in the Army can be misguiding when the results from different hospitals are compared. Taken at its face value, medical downgrading might signify that non-operative treatment in E.M.S. hospitals is not so good as in Military hospitals and that Army surgeons are not so good as civilians. In fact, the difference was due to the administrative necessity of treating all the more serious and long-term cases in E.M.S. hospitals.

D. Sulpha Drugs and Penicillin for the Treatment of Gonorrhoea (Part V).

Four issues call for further comment in this context.

1. The sooner the effectiveness of the treatment can be established, the sooner the effectiveness of the treatment can be estimated. Immediate criteria of relapse are extremely difficult to find even in such an acute infection as Gonorrhoea; and if delayed criteria are used there is always the possibility of confusing relapses with re-infections.

2. Relapses are not always clear-cut. Re-infection with respect to Gonorrhoea is easy, and occurs frequently as shown in the following table recast from figures given by Paley and Truelove (1946):

	No. of Cases	Percentage with 2 or more infections with Gonorrhoea.	Percentage with 3 or more infections with Gonorrhoea.
First Group	254	26.8	6.7
Second Group	161	20.5	5.6
TOTAL:	415	24.3	6.3

Relapses dealt with in Part V were in fact clear-cut, there being either a sufficient lapse of time to establish cure from a previous attack or a recent risk, acknowledged by the patient, of re-infection.

3. What constitutes a composite treatment for the purpose of assessing the efficacy of a single method is not a purely logical issue. It is one which calls for first-hand knowledge of current medical practice. In this instance, only treatments given simultaneously need be considered as composite. If given successively, we may assume that the first treatment failed.

4. Duration of stay in hospital or other medical unit, being subject to changing administrative arrangements, is rarely if ever a safe criterion in post hoc studies. In a civilian population it might be legitimate if applied in a single centre where it was possible to use alternate patients for experimental and control groups.

E. Penicillin Treatment of Impetigo (Part VI).

The issue of paramount interest raised by this enquiry concerns the delimitation of a satisfactory control group; and the conclusion which emerges is one which received in Part VI more honour in the breach than in the observance. In comparing different methods of treatment, it is advisable to compare new methods with a single method of known efficacy. Here two single methods of treatment of previously unmeasured efficacy were contrasted with a standard comprising a group of different methods whose effects were regarded as being those of a single method. There was, therefore,

an all too flimsy basis of comparison, since the standard embodied the mean of some possibly good therapies and some certainly bad ones. All that the trial conclusively proved was that Penicillin is certainly more efficacious than mere inaction.

The Impetigo trial like others emphasizes that uniformity of policy with respect to documentation and procedure is essential. The results might have been seriously biased by differences with respect to administrative measures current in the different hospitals, if no cognizance had been taken of them. Because Impetigo is an infective disease liable to occur in poor living conditions and with poor food, homogeneity of socio-economic level is an important prerequisite of satisfactory test subjects. The possible bias of nutritional factors and different environmental conditions was avoidable in this case because the sample was an Army population under treatment during the same period and while still resident in Britain.

BIBLIOGRAPHY

BURN, J. H., 1937. Biological Standardisation. Oxford University Press.

CLARK, A. J., 1940. Applied Pharmacology. Churchill, London.

De LIND van WIJNGAARD, 1926. Arch. exp. Path. Pharmak. 113:40.

LANDGREBE, F. W., 1939. J. Exper. Biol., 16:89.

MARKS, 1926. Publications of the League of Nations Health Organization. C.H. 398.

PADGETT, P., 1940. Am. J. Syph. Gonor., and Ven. Dis., 24:692.

PALEY, R.G. and TRUELOVE, S.C., 1946. Psychiatric Breakdown in the Syphilitic Population. Bull. Army Health Statistics, 2nd Series, 6th Issue.

SHAPIRO, and ZWARENSTEIN, 1933. J. Exper. Biol., 10.

SOLIMAN, T., 1942. A Manual of Pharmacolocy. Saunders, Philadelphia.

WHYTE, L., 1946. The Science and Art of Medicine. Inaugural Address. Cambridge University Press.